

University of Groningen

Mapping the dimensions of linguistic distance

Feleke, Tekabe Legesse ; Gooskens, Charlotte; Rabanus, Stefan

Published in:
Lingua

DOI:
[10.1016/j.lingua.2020.102893](https://doi.org/10.1016/j.lingua.2020.102893)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Feleke, T. L., Gooskens, C., & Rabanus, S. (2020). Mapping the dimensions of linguistic distance: A study on South Ethiosemitic languages. *Lingua*, 243, [102893]. <https://doi.org/10.1016/j.lingua.2020.102893>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Mapping the dimensions of linguistic distance: A study on South Ethiosemitic languages

Tekabe Legesse Feleke ^{a,*}, Charlotte Gooskens ^b, Stefan Rabanus ^a

^a Department of Linguistics, Verona University, Italy

^b Department of Linguistics, University of Groningen, The Netherlands

Received 3 January 2020; received in revised form 21 April 2020; accepted 22 April 2020

Available online



Abstract

We measured the distance among selected South Ethiosemitic languages from three dimensions: structural, functional and perceptual. The main objectives of our study was to determine the relationship among these three dimensions of linguistic distances, to re-examine previous classifications of the languages, and to measure the degree of mutual intelligibility among the language varieties. We determined the structural distance by computing the lexical and phonetic differences. The phonetic distance was computed using the Levenshtein algorithm. A Word Categorization test was adopted from [Tang and van Heuven \(2009\)](#) to measure the functional distance or the degree of intelligibility. A self-rating test, based on the recordings of 'The North Wind and the Sun' was administered to measure the perceptual distance among the languages. Then we performed a cluster analysis using Gabmap. Multidimensional scaling was employed for the cluster validation. The results of our study show that the selected South Ethiosemitic languages can be classified into five groups: {Chaha, Ezha, Gura, Gumer}, {Endegagn, Inor}, {Muher, Mesqan}, {Kistane} and {Silt'e}. Moreover, the language taxonomies obtained from the measures of the three dimensions of distance are very similar, and they are generally comparable to the classifications previously proposed by historical linguists. There is also a very strong correlation among the three dimensions of distance. Furthermore, the Word Categorization test results show that the majority of these languages are mutually intelligible with the exception of Silt'e.

© 2020 Elsevier B.V. All rights reserved.

Keywords: Ethiosemitic languages; Linguistic distance; Combined approach

1. Introduction

Issues of how to distinguish dialects from languages and how to quantify the resemblance between two or more language varieties have been the central concerns of dialectology. These two subjects are often addressed by measuring the distance between two or more language varieties. As a general principle, the more two languages are structurally similar (i.e., phonetically, morphologically, lexically or syntactically), the more they are related to each other; if they are similar enough, they are dialects of the same language. However, distinguishing dialects from languages is more complex because in most cases non-linguistic variables (social, cultural and political) have roles to play. This means that computing linguistic distance solely based on the structural similarity between languages may not always be sufficient to determine whether two language varieties should be considered as dialects of a language or two different languages. In

* Corresponding author.

E-mail addresses: tekabelegesse.felege@univr.it (T.L. Feleke), c.s.gooskens@rug.nl (C. Gooskens), stefan.rabanus@univr.it (S. Rabanus).

addition to the influences of the non-linguistic variables, there are inherent limitations of the structure-based traditional approach. The structural approach is often criticized for having two drawbacks. First, measuring the linguistic distance requires quantifying the distance among the language varieties. However, languages differ on several dimensions (phonology, phonetics, morphology and syntax) and identifying the level to be measured is a major challenge (Gooskens, 2018, p. 206; Heeringa et al., 2006, p. 51; Tang and van Heuven, 2007, p. 223; Tang and van Heuven, 2009, p. 710). Second, even if all the levels could be measured, determining the relative contribution of each level, and combining the differences into a single mathematical measurement, is another challenge (Chiswick and Miller, 2005, p. 1).

Previous studies of dialectology have generally followed two research perspectives to address the aforementioned limitations. On the one hand, there has been a successful move in terms of shifting from measuring linguistic distance solely based on purposefully selected specific linguistic features to measuring distance based on large aggregate data (Goebel, 2010; Nerbonne and Heeringa, 2001; Nerbonne et al., 2011; Prokić et al., 2013). On the other hand, different methods that take into account the non-linguistic variables, for example, the perception and the knowledge of non-linguists, have been developed in recent decades to circumvent the limitations of the structure-based approach (e.g., Preston, 2010). In this regard, the use of intelligibility as a means of measuring linguistic distance and recent advances in folk linguistics have made important contributions. As a part of these endeavors, different methods of measuring intelligibility have been developed (see Casad, 1987; Gooskens, 2013; Menuta, 2013, pp. 57–58; Voegelin and Harris, 1951).

There have also been various methods of measuring linguistic distance from perceptual perspectives. The perception-based approaches vary in the following ways. Some of them examine the perception of the speakers based on carefully selected language inputs, such as recorded stories (e.g., Beijering et al., 2008); some others measure the overall perception of the speakers without focusing on a specific language input, for example, by asking in which nearby area a similar language is spoken (e.g., Bucholtz et al., 2007; Pearce, 2009; Tamasi, 2003; Montgomery, 2007; Preston, 1996). Moreover, some recent studies have focused on examining the perception of non-linguists regarding specific sound features, such as the features of vowels or consonants (e.g., Labov, 2001; Plichta and Preston, 2005; Niedzielski, 1999). Therefore, dialectologists have taken different paths in attempting to better quantify the distance among related languages, leading to a substantial increase in methods for measuring linguistic distance. These methods can be subsumed into three broad categories: structure-based (based on phonetic, lexical or grammatical similarity), intelligibility-based (based on inherent and acquired intelligibility) and perception-based (based on the perception of native speakers). Previous studies measured linguistic distance either from one or from the combinations of these three perspectives (Casad, 1987, pp. 89–98; Gooskens, 2018, p. 196; Grimes, 1990; Tang and van Heuven, 2009, p. 710; Tang and van Heuven, 2007, p. 223; Tang and Van Heuven). As noted by Gooskens (2018), the degree of correlation among the linguistic distances measured from each of these perspectives is a concern that merits further exploration.

In the present study we further investigate this matter. For the sake of expediency, we use functional distance and intelligibility with slight differences in meaning. We adopt the common definition of intelligibility, which is the degree to which speech variety A is understood by the speakers of speech variety B on the basis of the linguistic similarity between A and B (Gut, 1980, p. 57). We define functional distance as the degree of difference between language A and language B on the basis of the speakers' degree of understanding. This distinction is important for some logical reasons. First, in literature, very often a distinction is made between inherent intelligibility and acquired intelligibility (Gooskens and Heuven, 2018, p. 2; Gut, 1980, p. 57). In some cases, only inherent intelligibility is considered as mutual intelligibility (e.g., Gut, 1980; Tang and van Heuven, 2009). We use functional distance to refer to a linguistic distance that is measured using either inherent intelligibility or acquired intelligibility tests or both. Second, both inherent intelligibility and acquired intelligibility are parts of actual communication – which is the main function of the language. Hence, functional distance (function-based distance) can best describe all distances measured from this perspective. More importantly, by using functional distance, we make a distinction between intelligibility, which is measured based on the actual performance, and perceived intelligibility, which is measured based on the subjective judgement of the native speakers. Based on these considerations, we classify methods of measuring linguistic distance in general as structure-based, function-based and perception-based. The distances that are determined using these methods are therefore considered as structural, functional and perceptual distances respectively.

By examining these three distances, we contribute to one of the continuing debates in dialectology, that is, to what extent these dimensions of distance correlate. In previous works, there have been doubts, for example, about the reliability of the awareness of non-linguists when measuring linguistic distance (Goeman, 1999, p. 141). The correlation between intelligibility and degree of linguistic similarity has also been the concern of several recent studies (Gooskens, 2018; Gooskens and Heuven, 2018; Gooskens et al., 2010). The present study partly addresses these concerns, and examines them in the context of Ethiosemitic languages. In addition to examining the relationship among different perspectives of measuring linguistic distance, we also aim to determine the distance and degree of intelligibility among selected South Ethiosemitic languages – Chaha, Inor, Ezha, Endegagn, Gura, Gumer, Mesqan, Muher, Kistane and Silt'e. These languages were selected on the basis of two parameters: the number of speakers of each variety and the

language sub-family the varieties belong to. Given that we wanted to include a high number of participants, language varieties with a relatively high number of speakers were selected based on the Ethiopian National Census Report (ENSA, 2007). We took the numbers of residents of the ethnically defined districts (e.g., Ezha Wereda, total population 84,905) as the numbers of speakers of the varieties. We also sought to include at least one language from each of the five so-called Gurage varieties: Kistane (North Gurage), Muher and Mesqan (West Gurage), Silt'e (East Gurage), Endegagn and Inor (Peripheral West Gurage) and Gura, Gumer, Chaha and Ezha (Central West Gurage).

2. Ethiosemitic languages

Ethiosemitic languages are Semitic varieties that are spoken in Ethiopia and Eritrea. Scholars have proposed different classifications of these languages. The present study largely relies on the classification of Hetzron (1972), which is considered as being most complete in terms of the number of languages included. Ethiosemitic languages are divided into North and South Ethiosemitic. The North Ethiosemitic languages consist of Tigre, Tigrigna and Ge'ez (see Demeke, 2001, p. 2; Hetzron, 1972, p. 3). The South Ethiosemitic languages include several languages (see Fig. 1). Languages classified under the 'Outer South' and 'Eastern' branches are traditionally called Gurage languages. According to Demeke (2001, p. 61), Fleming (1968, p. 354) and Faber (1997, pp. 3–4), there is no clear genealogical relationship among all of the Gurage varieties that constitute a large number of the South Ethiosemitic languages. For instance, Silt'e is closer to Harari than to the rest of the Gurage languages. Furthermore, Kistane is closer to Gafat than to other Gurage languages. There is also a controversy about the position of Mesqan. Hetzron (1972) and Hetzron (1977) classified it under West Gurage while other scholars, such as Demeke (2001), classified it under North Gurage. Moreover, Muher does not have a settled position in the classification of Ethiosemitic languages. While Hetzron (1972) classified it under the tt-Group, Demeke (2001) placed it under Central West Gurage. Neither of the studies provided a sufficient description for their classification. Lack of detailed evidence, combined with other factors, such as a long history of contact among Ethiosemitic and other neighboring Afro-asiatic languages, compelled previous studies to make often unsatisfactory conclusions regarding the

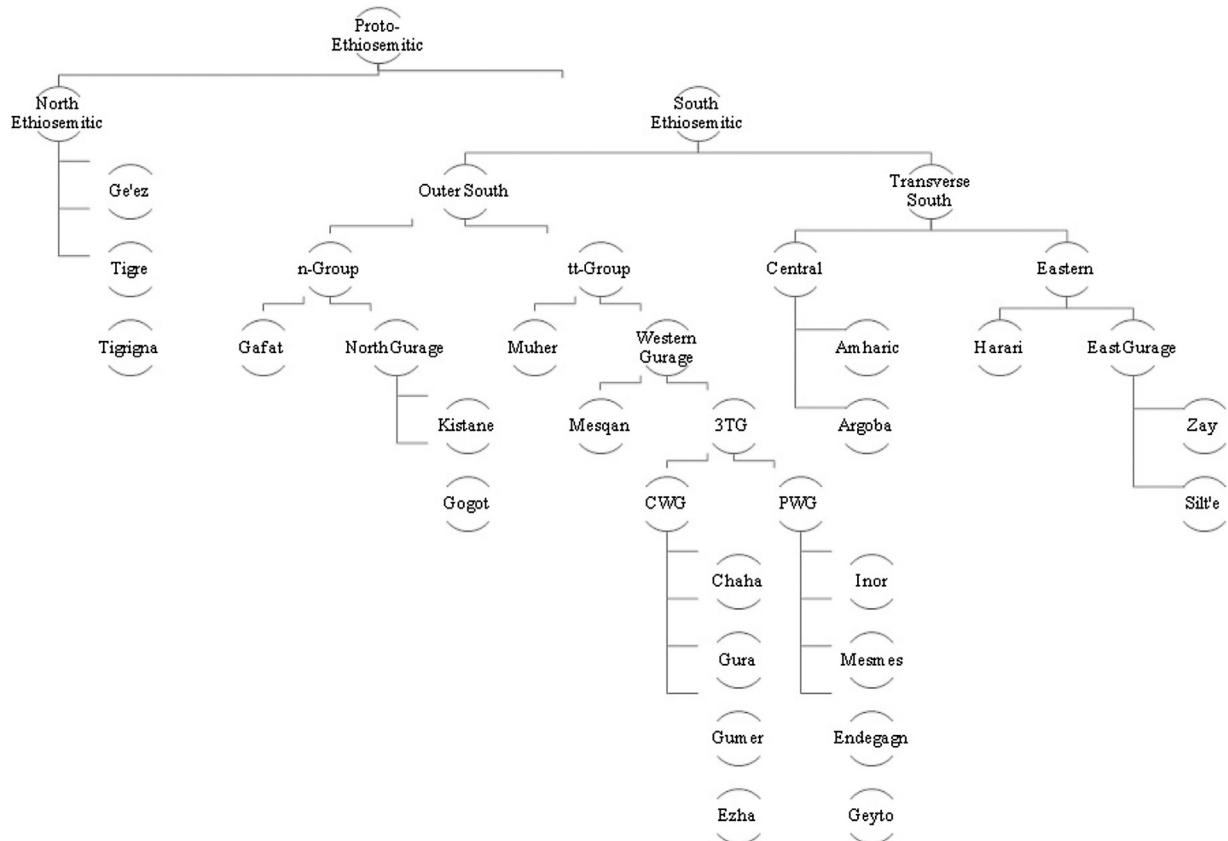


Fig. 1. Classification of Ethiosemitic languages (Hetzron, 1972). Hetzron (1972) considered Wolane as a dialect of Silt'e.

origin and the classification of the languages (Goldenberg, 1977, p. 462). So far, there is no single clear proposal regarding the origin and the classification of Ethiosemitic languages (Demeke, 2001, p. 61; Hetzron and Bender, 1976, p. 5; Hudson, 2000, pp. 75–76; Goldenberg, 1977, p. 461; Leslau, 1969, p. 97; Leslau, 1992, p. 12).

There are no studies concerning phonetic and perceptual distances among Ethiosemitic languages. However, there are studies on lexical comparisons. For instance, Bender et al. (1972) examined 12 Ethiosemitic languages using a 98-words list from Swadesh (1955). Inor, Chaha, Mesqan and Kistane are among the languages included in this study. According to Bender et al. (1972), none of these languages share more than 80% of cognates. Bender (1971) also made a lexical comparison among several Ethiopian languages. The study adopted a 98-word list from Swadesh (1950) and found out that none of the compared Ethiosemitic languages (Amharic, Argoba, Harari, zay, Wolane, Inor, Chaha, Mesqan, Kistane, Gyeto, Ge'ez and Tigrigna) shared more than 80% of cognates. In the same manner, Hudson (2013) investigated the lexical similarity among 14 Ethiosemitic languages based on a 250-word list. Silt'e, Inor, Chaha, Muher, Mesqan and Kistane are among the languages investigated by this study. The study reported more than 80% shared cognates between Inor and Mesqan, Inor and Muher, Inor and Chaha, Chaha and Muher, Muher and Mesqan, and Mesqan and Kistane. Likewise, using a list of 255 words, Menuta (2015) examined the lexical similarity among six South Ethiosemitic languages: Kistane, Chaha, Inor, Mesqan, Muher and Wolane. The study reported more than 80% of cognates were shared between Chaha and Inor, Chaha and Mesqan, Chaha and Muher, Mesqan and Chaha, and Mesqan and Muher.

The degree of intelligibility among many of the languages has not been investigated either. To the best of our knowledge, there are three studies that have investigated the degree of intelligibility among South Ethiosemitic languages: Gutt (1980), Ahland (2003) and Menuta (2015). Gutt (1980) examined the intelligibility among six South Ethiosemitic language varieties, Silt'e, Kistane, Chaha, Inor, Mesqan and Amharic, using an oral comprehension task. The results of the study indicate that, based on the 80% intelligibility threshold, only Silt'e and Mesqan are intelligible. In the same manner, Ahland (2003) determined intelligibility among eleven Gurage varieties using oral comprehension questions. According to Ahland (2003), based on an 80% intelligibility threshold, Chaha is intelligible to Ezha, Muher and Gumer; Ezha is intelligible to Gumer; Inor is intelligible to Endegagn; Gumer is intelligible to Ezha and Endegagn; Endegagn is intelligible to Inor; Mesqan is intelligible to Chaha, Ezha and Muher. Menuta (2015) also investigated intelligibility among six Gurage varieties (Kistane, Mesqan, Inor, Chaha, Muher and Wolane). In this study, different tests were used to measure intelligibility: word recognition (words in different parts of sentences were recognized by the respondents), sentence repetition (the informants listened to various sentences and wrote down exactly what they heard), sentence verification (the informants judged sentences that are habitually true by saying 'true' or 'false'), instruction (the respondents performed certain actions based on given instructions) and comprehension questions. Based on the 80% intelligibility threshold, this study reported intelligibility between Chaha and Inor, Chaha and Mesqan, Inor and Mesqan, Mesqan and Kistane, Muher and Chaha, and Muher and Mesqan.

With regard to the geographical distribution of the languages, Ethiosemitic languages in general are spoken in the north, central, east and southwest of Ethiopia. The 10 languages we investigated in the present study are spoken in the southwest part of Ethiopia (see Fig. 2), around 160 km from Addis Ababa, the capital. This small area is sometimes called the Gurage area. It is one of the most linguistically diverse areas in Ethiopia. More than 12 Ethiosemitic varieties are spoken in this area. We adopted the term 'Gurage language area' and 'Gurage languages' from earlier works (e.g., Leslau, 1979). However, it is important to mention here that the so-called Gurage languages do not refer to a single genetically attested unit (Hetzron, 1972, p. 119; Meyer, 2011, p. 1221) and, more importantly, some of the speakers of these varieties do not consider themselves as Gurage (Meyer, 2011, p. 1223). Silt'e is taught at elementary level in Silt'e Zone while the remaining varieties are just spoken languages.

Given that there have been debates concerning both the methods of dialectology and the classification of Ethiosemitic languages, the present study aims to address two general objectives. The first one is methodological, that is, to what extent the methods of measuring linguistic distance are related. There are two specific objectives related to the methods: (a) determining to what extent structural, functional and perceptual distances correlate; (b) examining the possibility of substitutability among the three dimensions of distance. By addressing these objectives, we illustrate the link among various methods of measuring linguistic distance. We expect strong correlations among the three dimensions of distance based on previous studies (e.g., Beijering et al., 2008; Casad, 1987; Grimes, 1990; Gooskens and Heeringa, 2004; Tang and van Heuven, 2007; Tang and van Heuven, 2009; Van Bezooijen and Gooskens, 2007).

The second general objective is determining the linguistic distance among the selected South Ethiosemitic language varieties. We aim to address four specific objectives related to the Ethiosemitic language varieties: (a) determining the distance among the selected language varieties; (b) classifying the languages using the data obtained from the three dimensions of distance; (c) examining to what extent the taxonomies obtained from structural, functional and perceptual distance measures are similar to the classifications previously proposed by historical linguists; (d) determining the degree of intelligibility among the language varieties. Based on Hudson (2013) and Menuta (2015), we expect very close lexical similarity between Chaha and Mesqan, Chaha and Inor, Mesqan and Muher, and Mesqan and Kistane. Furthermore, we

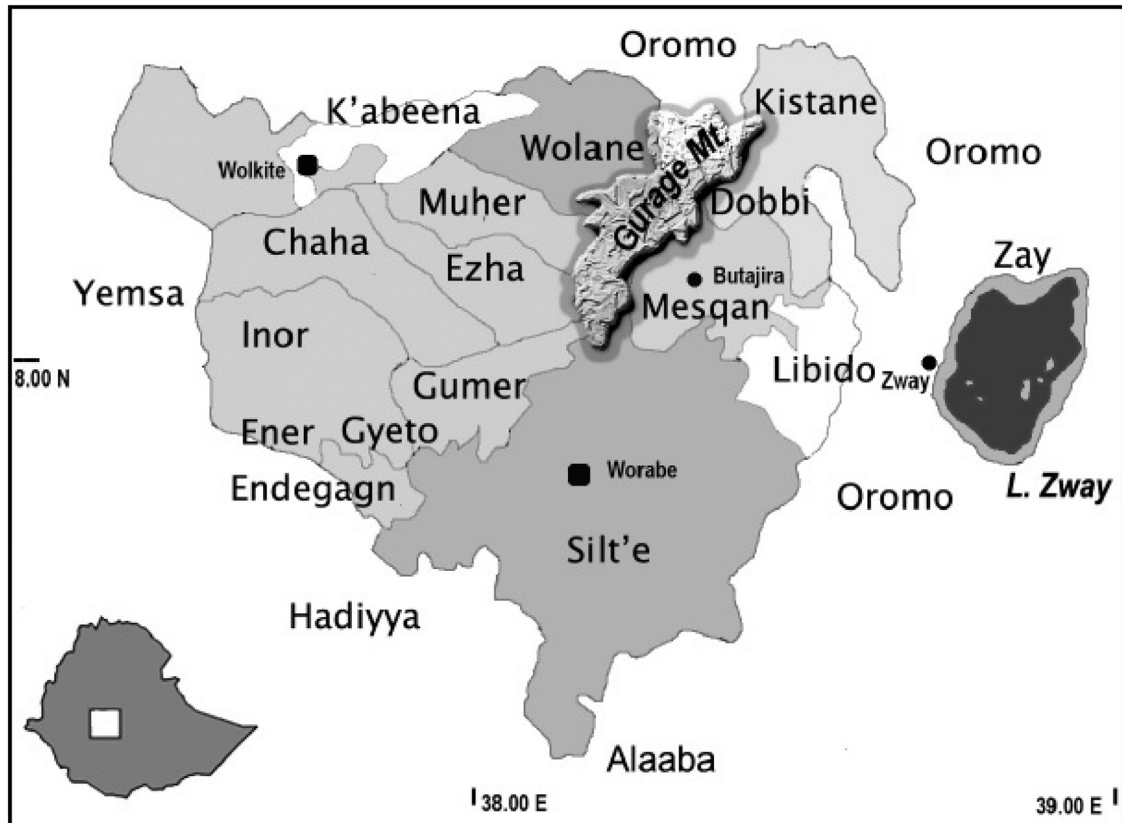


Fig. 2. Gurage language area, from Meyer (2014). The dark color represents Zway lake around the Gurage area. Oromo, Alaaba, Hadiyya and K'abeena are Cushitic languages spoken around the Gurage area. Yemsa is an Omotic language. Gura is spoken in Chaha area while Wolane is spoken in Silt'e area, and the geographical boundary of Gura and Wolane has not been identified yet.

expect close similarity between taxonomies obtained from the three distance measures, as well as from the classifications previously provided by historical linguists, based on Tang and van Heuven (2009). Based on Ahland (2003) and Menuta (2015), we further expect intelligibility between Chaha and Ezha, Chaha and Muher, Chaha and Gumer, Mesqan and Chaha, and Mesqan and Ezha.

3. Methods

This section presents the methods employed to address the objectives presented in Section 1. First, the description of the research assistants and experimental subjects is presented. This will be followed by the methods and procedures used to measure structural distance. Then, methods used to determine functional and perceptual distance are explained. This is followed by a presentation of the methods of clustering and cluster validation techniques.

3.1. Research assistants and experimental subjects

In this study, the term 'research assistants' and 'experimental subjects' are used with different meanings. Research assistants were school teachers who participated in selecting test-takers, preparing materials, such as translating texts, and reading translated texts during the recordings. 'Experimental subjects' refer to student who completed the tests designed to measure functional and perceptual distances. The procedures used to select both the research assistants and the experimental subjects are presented as follows.

3.1.1. Research assistants

Research assistants refer to selected secondary school teachers (minimum bachelor degree holders). They were selected from 10 schools in nine districts in the Gurage and Silt'e areas: eight districts in Gurage Zone and one district in Silt'e Zone (Chaha and Gura are spoken in Chaha district). From each school, three teachers who spoke the variety of that particular area as a native language, were selected. In other words, a total of thirty teachers were recruited from the 10 schools in the 10 districts. The teachers were selected using two screening steps. For the initial screening, a call for participation in the form of printed leaflets, was distributed in the schools. The leaflets explained a few language requirements, such as being native speakers of the local variety and having lifelong residence in the language area. There were many schools in some of the districts. Except for Mesqan and Gura, a school in the administrative town of each district was selected. Regarding Mesqan, the administrative town is Butajira. Since the residents of Butajira are largely Amharic speakers and Mesqan is not frequently used, a school outside the Butajira town was selected. Gura is spoken in Chaha district. Regarding Gura, speakers from around Gura Megenase (a suburban area of Endebir, a town in Chaha) were considered.

On the leaflet the contact information of the principal investigator was included so that any interested teacher could easily get in touch with the researcher if the requirements were fulfilled. The call for participation was posted on the notice boards of all the secondary schools in the districts of interest. Among the teachers who responded to the call for participation, three were selected from each language area. This second screening was conducted using semi-structured interviews. The interviews focused on issues such as the teachers' home language situation, the amount of exposure to the neighboring varieties, and language conditions in earlier workplaces (whether they regularly used native language in the workplace). Based on these parameters, teachers who were native speakers of the local variety, and who used the language both in schools and at home, were recruited. The interviews took place in the schools of the respective teachers. They received a small amount of payment (300 birr) for their services.

3.1.2. Experimental subjects

The subjects were selected by the research assistants. Thirty students were recruited from each school, 300 in total. The students in all the grade levels in the secondary schools (from grade 9 to 12) were considered in order to include as many students as possible. Similar to the selection of the research assistants, the students were selected in a two-step screening process. First, all students who are native speakers of the local variety were requested to complete a registration form prepared for this purpose. The registration was completed by the research assistants. Once the native speakers of a local variety were identified, they were submitted to the second screening. Questionnaires were employed for the second screening. The questionnaires contained items about the students' first and second language background, family language conditions, personal information and their contact with speakers of other neighboring language varieties. The questionnaires were prepared in Amharic because all secondary school students in the study areas were able to read and write Amharic. Indeed, Amharic is both the language of schooling and of workplaces in the study areas, except in Silt'e Zone where Silt'e is taught in elementary schools. The questionnaires were coded for each school and for each study area so that they could be easily identified during the analysis. All the items (questions) in the questionnaires were close-ended to maximize the accuracy of the responses and to take into account the age and education levels of the students.

Then, based on the information obtained through the questionnaires, 300 participants (30 from each variety), who are the native speakers of the varieties of interest, were selected. Furthermore, based on the data that were obtained from the questionnaire, it was assured that the participants had lived their whole life in the area where their variety is spoken and that their parents are native speakers of the variety under investigation. Whenever the eligible students that met the requirement exceeded 30 for each variety, the equal proportion of sex (15 males and 15 females) was used as an additional parameter. Whenever there were too many eligible candidates, 15 male and 15 female students were randomly selected. Prior to data collection, permission was obtained from both Gurage and Silt'e Cultural and Tourism Bureaus, and from the administration of each school. Not all the selected participants attended the tests. Given that the Word Categorization and perception tests were administered at different times, in some of the language sites, the number of participants who completed the Word Categorization test and the perception test was not exactly the same. In total, 285 participants completed the Word Categorization test. Among these, 171 were male and 114 were female. Moreover, 289 participants took part in the perception test, among which 171 were male and 118 were female.

3.2. Determining structural distance

Structural distance was measured from two perspectives: lexical and phonetic. Words for the structural distance measure were randomly collected from different sources: from a list of words gathered for the Word Categorization test, from a fable entitled 'The North wind and the Sun' (all the words in the story were included), as well as other published materials. Hence, a total of 240 words were compared to determine the two distances (see [Appendix B.1](#)). The principal

investigator orally presented the list of words to the research assistants from each language area, and asked them to provide (orally) equivalents of their native language for each of the words. Then, the principal investigator phonetically transcribed the equivalents provided by the assistants. In case of disagreements during translation, the assistants were told to resolve by majority rule (2/3).

3.2.1. Lexical distance

The lexical distances among the 10 selected language varieties were determined by computing the percentage of non-cognates of the total lexical items within pairs of varieties. Non-cognates are words that share meaning, but have different forms. The corpus of the lexical distance measurement is constituted by the words indicated in 3.2. The shared cognates were determined based on two parameters: similarity of roots and similarity of meaning between the corresponding pairs of words. These parameters were employed in a two-step process of cognate identification. First, the principal investigator identified pairs of words that share a common root based on the form (phonological) similarity between the corresponding words. In almost all Semitic languages, sequences of consonants form the basic word meaning (root). Hence, root similarity was considered as a core parameter, e.g., Amharic *b?re* 'ox', Endegagn *bawra* 'ox', Chaha *bora* 'ox'. Then, the similarity in meaning among the pairs of words that share the same root was confirmed by the principal investigator and the research assistants, who are native speakers of the varieties. Once the cognate and non-cognate words in pairs of all varieties were identified, the percentage of non-cognate words was computed.

3.2.2. Phonetic distance

The output of the lexical distance measurement was used as an input for the phonetic distance measurement, that is, phonetic distance was measured only between cognates that were phonetically transcribed. Cognates that are shared at least by six of the 10 language varieties were considered for phonetic distance. The cognates were aligned, and the distance among them was computed using the Levenshtein algorithm, based on the number of phones which were inserted, deleted or substituted. The distance computation was made using the simplest cost assignment. The simplest cost assignment assigns equal cost (1 unit) to all the operations. Only the distance among the cognates was computed based on Kessler (1995, p. 5) since the difference among non-cognates is not phonetic. The Levenshtein distance among the cognates was computed using Gabmap (see Heeringa, 2004; Nerbonne et al., 2011). The following are sample Levenshtein (phonetic) distances between Kistane and Chaha based on a shared cognate 'cloud'. In this case, the Levenshtein distance is 2; substitution of [m] by [b] and [n] by [r]. This operation costs two units. This distance value is divided by the longest alignment, 6 in this case, to obtain the normalized distance. The normalized distance between Kistane and Chaha in this particular example is .33 (2/6) (Table 1).

3.3. Functional and perceptual distances

This section presents tests designed to measure functional and perceptual distances among the 10 language varieties.

3.3.1. Functional distance

The Word Categorization test was adopted from Tang and van Heuven (2009). This test was selected since it could be administered with minimal priming or learning effect, the major factor that probably influenced previous studies by Gutt (1980), Ahland (2003) and Menuta (2015).

Materials: The material selection and preparation procedures were quite similar to those of Tang and van Heuven (2009). The first step in the material preparation was determining 10 semantic categories to be used for the test. The semantic categories were general concepts such as plants, fruits, animals, furniture, etc. (see Appendix B.2). One of the parameters was the frequency of use of the semantic categories among the speakers of all varieties. For instance, some

Table 1
Phonetic distance, using Levenshtein Algorithm.

Kistane – Chaha 'cloud'					
d	a	m	?	n	a
d	a	b	?	r	a
		1		1	
Absolute				2	
Relative				0.33	

categories, such as musical instruments, are extremely culture-specific; as a result, they might not be common among all speakers. The second parameter was the possibility of a semantic category incorporating as many words as possible. This parameter was important because each semantic category must contain at least 10 words. First, the principal investigator selected the categories based on his intuition. The categories were later approved by the research assistants.

Similar parameters were used to determine words to be included under each semantic category. In this case, word frequency was computed since frequency could be one of the factors that determine the comprehension of the words. It was not possible to compute directly the frequency of the lexical items to be categorized under each semantic category. This was because none of the Ethiosemitic varieties under investigation has its own structured corpus. Also many of them do not have online oral and written documents which could be used as input to create a corpus. The only language in the area with a sufficient amount of easily available language data is Amharic. Hence, an Amharic language corpus containing about 100,000 written words, was created using AntConc software (Anthony, 2004), and this corpus was used to estimate the frequency of each lexical item. All the sources of the data were written texts such as newspapers, magazines, academic articles and social media texts. In the corpus, texts of different genres (politics, economics, agriculture, culture, sport, science, etc.) were included to make the corpus as representative as possible. Using this corpus, words that have a relatively high frequency were selected.

Using these procedures, 10 semantic categories, each containing the 10 most frequent words, were identified (see Appendix B.2). After the identification of the words and the semantic categories, the words under each semantic category were translated by the research assistants from Amharic to the 10 varieties. The translators were told to solve the disagreements by majority vote (2/3) whenever there was a disagreement among them. After the translation, each translator pronounced the translated words, 100 words for each variety, for sound recording with Adobe Audition running on a personal computer. Then, the three translators from each variety were asked to rate their three recordings of 100 words on a Likert scale that ranged from 0 (not natural) to 5 (natural). Finally, among the three recordings, the one with the highest rating score was selected for the intelligibility test.

Procedure: In the Word Categorization test, the participants' recognition capability was tested through semantic-multiple choice categorization. In the test, the listeners indicated which of the 10 given semantic categories a spoken word belongs to. For instance, the respondents heard 'banana' and were asked to categorize this word under one of the 10 semantic categories ('fruits' in this case). The assumption here was that the correct categorization is achieved only if the listeners correctly recognize the target words. As there were 10 semantic categories for each word, the probability of categorizing the words by chance is very small (10%). In the process of developing this test, the primary activity was creating audio input so that the listeners do not hear the same word in the same variety more than once. In other words, the priming effect due to the repetition of similar input should be controlled for. Similar to Tang and van Heuven (2009), the Latin Square system was used for this purpose. Different data files (CDs) were created using the following procedures.

As indicated above, in the Word Categorization test, listeners must not hear the same word more than once. A word which is heard twice or more has more possibility of being recognized than a word which is heard only once – the priming effect. In the present study, there were 10 semantic categories, each semantic category consisted of 10 lexical items, a total of 100 (10×10) words. Based on these words, different CDs were created. On the first CD, the selected list of 100 words were presented in a fixed random order (1–100) in such a way that every subsequent word is spoken in different variety. This is a default order. On the second CD (CD2), the words were presented in the same order except that the presentation begins with the variety in which no. 100 was spoken, then followed by varieties in which no. 1 to no. 99 were spoken. Due to this shift, every word in CD2 was spoken in a different variety as compared to CD1. The third CD begins with the variety in which no. 99 was spoken followed by the variety in which no. 100 was spoken, followed by varieties in which no. 1 to no. 98 were spoken. Through this rotation, a total of 10 CDs were created, each CD containing 100 words in 10 semantic categories.

One CD was administered for participants from each language area (see Fig. 3). The 100 words on a CD were divided into 10 tracks and each track was presented to a group consisting of three participants (every track was repeated three times) so that each member of the group classified the 10 same words into 10 semantic categories. Since there were 10 tracks on each CD, a total of 30 students listened to each of the CDs administered in each language area. These procedures meant that: (1) each listener experienced each word only once. (2) A listener from every language area heard each word in 10 different varieties; and (3) Every member of a group heard one-tenth (1/10) of the total lexical items. Fig. 3 below shows the procedure of the task. Tang and van Heuven (2009) used 7 s as the response time. In the present study, the time was increased to 10 s in order not to put the students under time pressure. Before the actual testing, there was a practice session. For this session, a separate practice CD containing 10 words and 10 semantic categories from additional material was prepared. Each participant practiced at least once before beginning the actual task. More than one practice session was allowed depending on the confidence and interest of a participant.

For each track of the CDs, there was an answer sheet. Each answer sheet had its own CD and track numbers (e.g., CD 1, Track 2) so that each participant received an answer sheet with a different code number. Tang and van Heuven (2009)

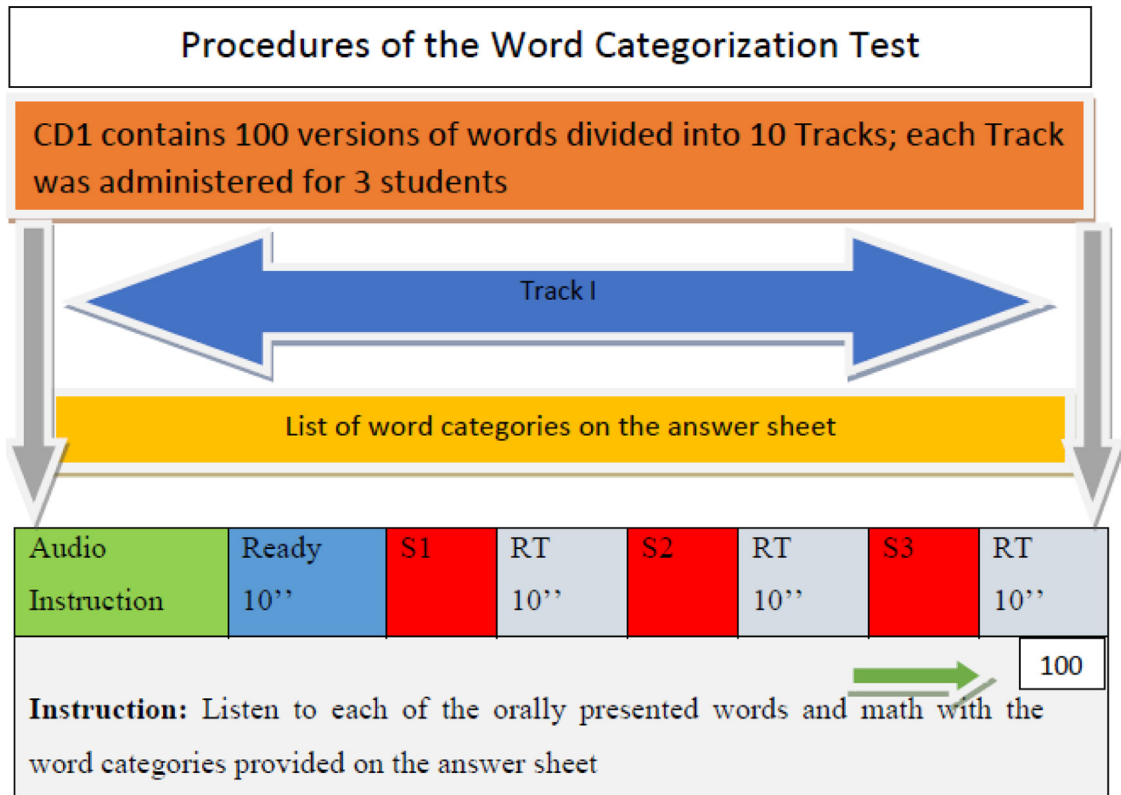


Fig. 3. Procedures of Word Categorization test.

provided the list of 10 semantic categories on the response sheet. The same method was used in the present study. After listening to the orally presented words, the participants responded by choosing the appropriate match from lists of categories provided on the response sheet. The test was administered in quiet classrooms in the selected schools. Each participant was tested individually in a separate session. The test was administered by the principal investigator and one of the research assistants. The intelligibility measure was the percentage of words correctly matched with the semantic categories provided.

3.4. Perceptual distance and attitude tests

This section presents procedures which were employed to determine perceptual distance and the attitudes of the speakers towards the test languages. Perceptual distance was measured from two perspectives: perceived similarity and perceived intelligibility. The presentation begins with the materials used for preparing the tests.

3.4.1. The materials

The fable 'The North Wind and the Sun' was used as input to determine the perceived intelligibility, the perceived similarity and the attitude of the speakers towards each other's variety. First, the fable was translated from English to each of the local varieties by the three research assistants recruited from each language area. In case of disagreements, the assistants were told to resolve using majority rule (2/3). A modified Amharic writing system was used for the translation. After the written translation, the translated version of each variety was orally presented by each of the three research assistants. The presentation of each translator was recorded using Adobe Audition running on a personal computer. Then, the three translators listened to each recording and rated the readings on a Likert scale that ranged from 1(not natural) to 5 (natural). Finally, among the three readings, the one which received the highest rating score was selected for the test. The recording was made in a silent room in each school. The recording process was administered by the principal investigator.

3.4.2. The tests and test procedures

The three types of tests: perceived intelligibility, perceived similarity and the attitude of the speakers were combined and administered at the same time using the same material. Each test was represented by one item (question) with its own rating scales. This means that the combined test contains three questions: one for perceived similarity; another for perceived intelligibility and the remaining one for language attitude. The three test items were presented simultaneously to minimize the effect of the participants' familiarity with the test material, that is, the test-takers answered the three questions after listening to each version of the recordings.

In order to minimize a response bias that might occur due to fatigue and familiarity with the test, the test items were arranged in three different orders; order A: (1) attitude test item, (2) perceived intelligibility test item, (3) perceived similarity test item; order B: (1) perceived intelligibility test item, (2) perceived similarity test item, (3) attitude test item; and order C: (1) perceived similarity test item, (2) attitude test item, (3) perceived intelligibility test item. Due to these arrangements, each test item appeared in three different orders. Before the test administration, the 30 speakers of each variety were randomly divided into three groups, each group containing 10 members. Then, the tests were administered in such a way that members of the same group received tests in the same order: the first group received order A, the second group order B and the third group order C. Administering tests of the same order for members of the same group was important to give the same instruction for all group members. The audio inputs were presented using a loudspeaker so that it would be possible for us to follow each response.

During the test, the test-takers listened to the recording of each variety and responded to the three successive questions. They responded by putting an 'X' on the Likert scale provided for each question. To measure the perceived intelligibility, the participants were asked to determine to what extent they understood the speaker in the recordings. After listening to each of the recordings, the test-takers indicated their judgment on the Likert scales that ranged from 0 ('do not understand at all') to 10 ('completely understand'). In the same manner, for perceived similarity, the respondents were asked to determine to what extent each of the presented recordings was similar to their own variety and to express their judgment using 11-point Likert scales that ranged from 0 ('not similar') to 10 ('completely similar'). With regard to the language attitude, the respondents were instructed to determine whether the language in which the story was presented was beautiful or not, and to provide their responses on 10-point Likert scales that ranged from 1 ('not beautiful') to 10 ('beautiful'). The recordings of the 10 language varieties were presented in different orders for the speakers of each variety to manage the impact of fatigue (respondents might be less conscientious on the last story). In other words, there were 10 different orders of the recordings, one order for the speakers of each language variety.

After the presentation of each recording, there was a response time of 3 minutes, 1 minute for each test item. For the sake of uniformity, the instruction was given in Amharic either by the principal investigator or by one of the research assistants. If there was a misunderstanding, further explanation was provided in the participants' native language. The recordings were presented using a personal computer attached to a loudspeaker. After listening to each recording, the listeners provided their responses by marking 'X' on the scale provided. For each recording, there was a separate answer sheet. In other words, each test-taker received 10 pages of response sheets, one page for each recording. This procedure was vital to make sure that the test-takers precisely matched each recording with the respective test items. The 30 selected students took part in the perceptual tests after they had taken part in the intelligibility test.

3.5. Clustering and cluster validation

After data collection, Gabmap was employed for the clustering and cluster validation. Gabmap is dialect classification and visualization software developed by linguists at the University of Groningen (see [Leinonen et al., 2016](#); [Nerbonne et al., 2011](#); [Snoek, 2014](#)). It provides several statistical alternatives to group similar languages together. Based on [Gooskens and Heeringa \(2004, p. 196\)](#), the Weighted average method was employed to classify the Gurage varieties. However, clustering is often tricky – a small variation in the data matrix could result in different groupings. Gabmap provides three clustering validation techniques – discrete clustering, fuzzy clustering and multidimensional scaling. In the present study, multidimensional scaling was used to make sure that the clusters created were valid and reliable (see [Nerbonne et al., 2011](#)). The results of fuzzy clustering are presented in the appendix (see [Appendix C.1](#)). Multidimensional scaling takes a distance matrix as an input and groups values that are similar. Gabmap provides multidimensional scaling in a two-dimensional space. The first dimension explains much of the variance in the distance matrix. The second dimension explains a large portion of the remaining variances.

4. Results

Various distance matrices were obtained from the structural, functional and perceptual distance measures. In Section 4.1, we report results of the classifications of the language varieties based on structural, functional and perceptual

distances. As indicated in Section 3, structural distance was measured using the phonetic and lexical differences. Functional distance was determined based on the respondents' scores on the Word Categorization test, perceptual distance was estimated based on the respondents' response to the self-rating perception test. The average of the upper and the lower halves of the distance matrix was considered as the distance between languages in both the functional and perceptual measures. Section 4.2 presents the results of the relationship among the three dimensions of distance. Section 4.3 presents the results of the Word Categorization test.

4.1. Classifications of the South Ethiiosemitic languages

In this section we present the classifications of the South Ethiiosemitic languages based on the measures of the three dimensions of distance. The classification results are supplemented by the results of multidimensional scaling.

4.1.1. Classification of the languages based on structural distance

Fig. 4(a) shows the multidimensional scaling plot of phonetic distance in two-dimensional space. The first dimension is indicated by a solid arrow and the second dimension by a dashed arrow. In Fig. 4(a), the first dimension shows that Chaha, Gura, Gumer and Ezha have low values while Kistane and Silt'e have the highest values. The values of other languages are between these two extremes. This dimension explains 52% ($r = .72$) of the variance in the distance matrix. The second dimension (dashed arrow) indicates that Endegagn has the lowest value while Mesqan and Muher have the highest values. The values of other varieties are between these two extremes. This dimension also explains 38% of the variance ($r = .62$). The two dimensions combined explain 90% of the variance in the matrix. Based on phonetic distance, the multidimensional scaling plot indicates six groups of language varieties: {Chaha, Gura, Gumer, Ezha}, {Muher, Mesqan}, {Endegagn}, {Inor}, {Silt'e} and {Kistane}. As can be seen from Fig. 4(a), Silt'e and Kistane are separate languages. Inor and Endegagn are also phonetically somehow different. Fig. 4(b) shows that map of the first dimension of the

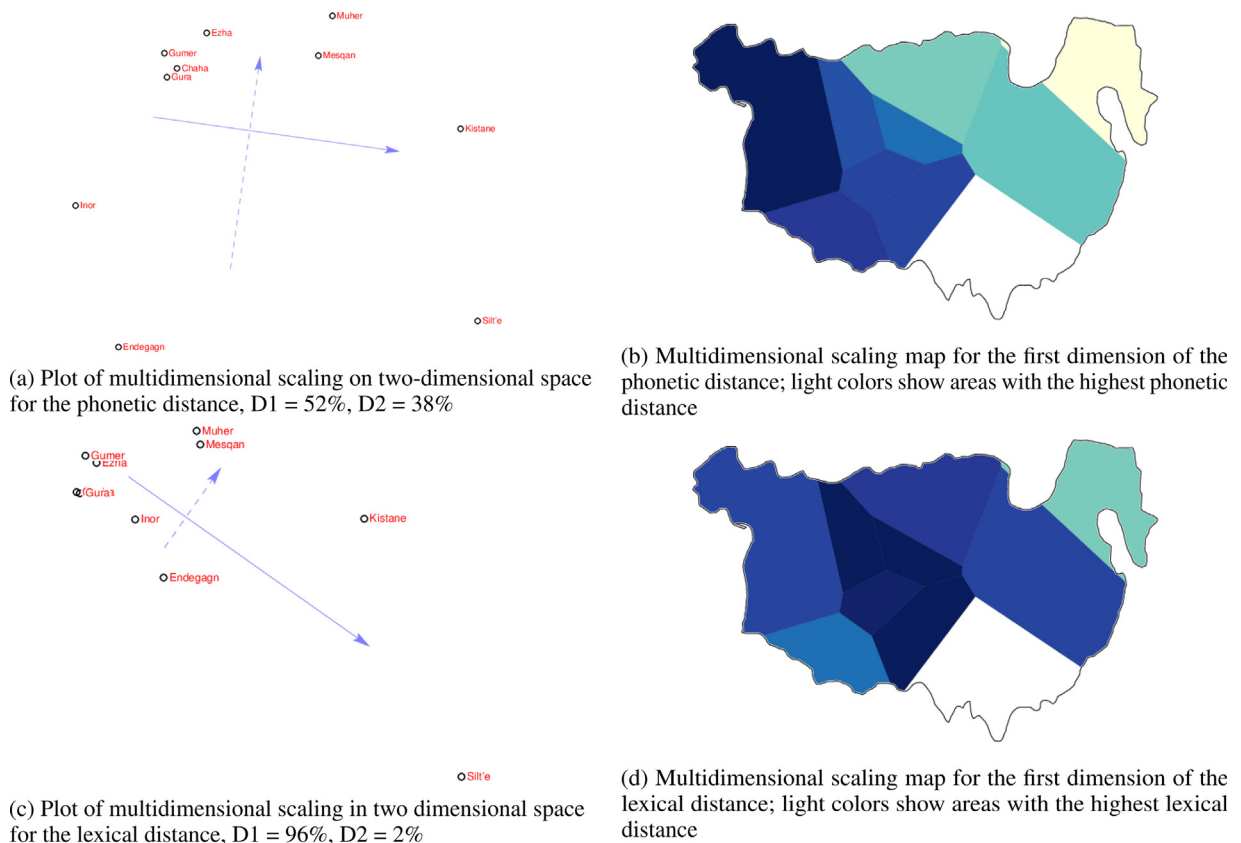


Fig. 4. Classification of South Ethiiosemitic languages based on structural distance.

multidimensional scaling for the phonetic distance. The light color shows the area with the highest linguistic distance, which is the Silt'e area.

The multidimensional scaling plot based on lexical distance is illustrated in Fig. 4(c). The first dimension (solid arrow) explains the majority of the variance, 96% ($r = .98$). As the figure illustrates, Gura, Gumer, Chaha, Ezha have the lowest values, and Silt'e has the highest value. The values of the other varieties are somewhere between these two extremes. The second dimension (dashed line) shows that Inor has the lowest values while Muher and Mesqan have the highest values. This dimension explains 2% ($r = .15$) of the variance in the data matrix. The two dimensions combined explain 98% of the variances. The multidimensional scaling plot of lexical distance shows five possible groupings of the language varieties: {Gumer, Gura, Ezha, Chaha} form a group. {Inor and Endegagn} also form a group. In the same manner, {Muher and Mesqan} form a group. However, {Kistane} and {Silt'e} are separate languages. Fig. 4(d) illustrates the map of the first dimension of multidimensional scaling for the lexical distance. The light color shows areas with the highest linguistic distance which is again the Silt'e area.

The dendrograms obtained from the distances are presented in Fig. 5(a) and (c). The two dendrograms illustrate the classification of the language varieties based on the phonetic and lexical distances, respectively. Fig. 5(b) and (d) illustrate the dialect maps of the language varieties based on phonetic and lexical distances respectively. As can be seen from Fig. 5(a), based on phonetic parameter, {Gura, Gumer, Ezha, Chaha} form a group. {Muher and Mesqan} are closely related. However, {Kistane} and {Silt'e} are separate languages. Likewise, {Endegagn} and {Inor} are separate languages. Fig. 5(b) also shows the geographical distribution of the six dialect areas. In general, the phonetic distance measure shows that the South Ethiopese languages are classified into six dialect areas.

Fig. 5(c) presents the dendrogram of the language varieties based on lexical distance. As can be seen from the figure, from a lexical point of view, {Gura, Gumer, Chaha and Ezha} form a group. {Endegagn and Inor} also form a group.

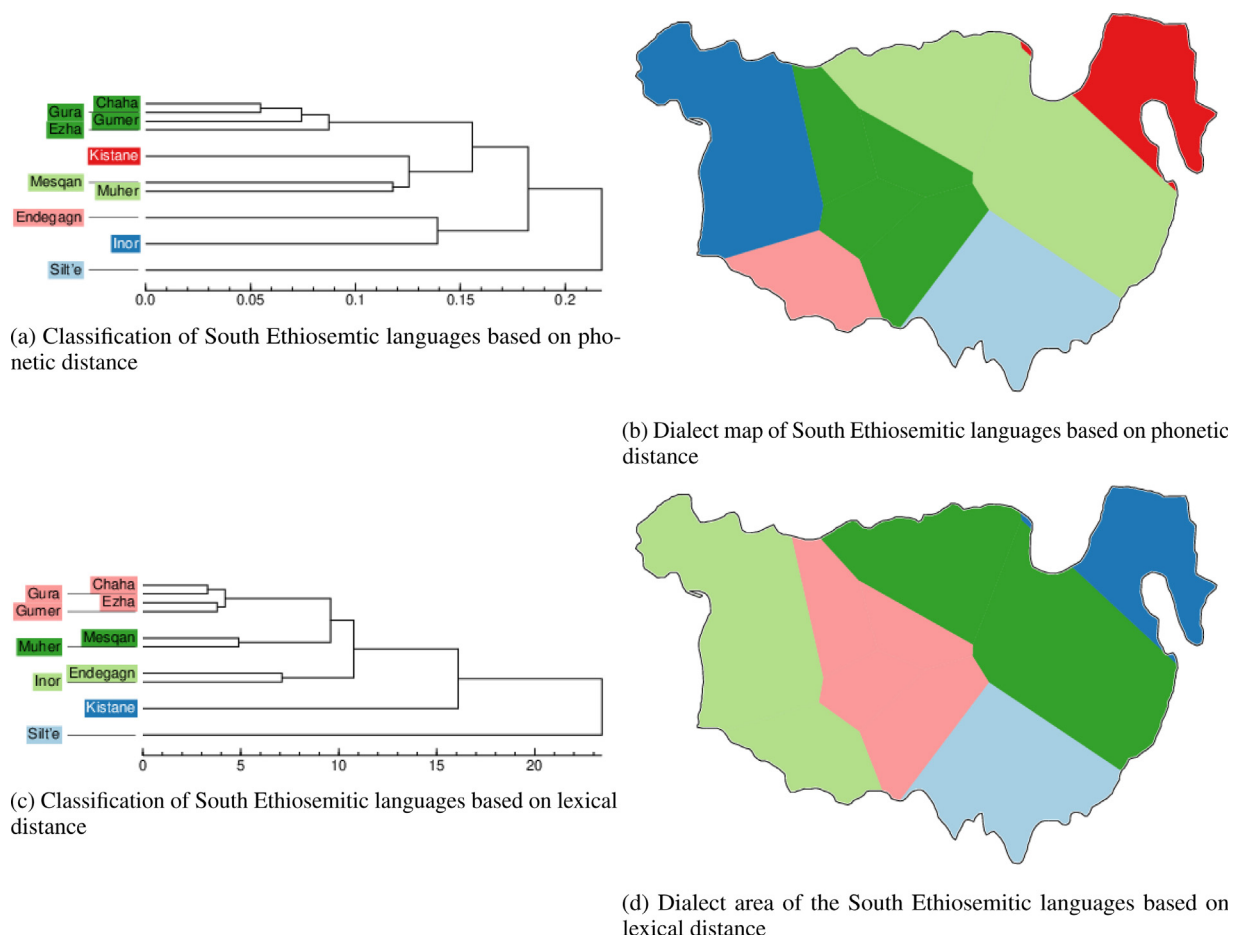


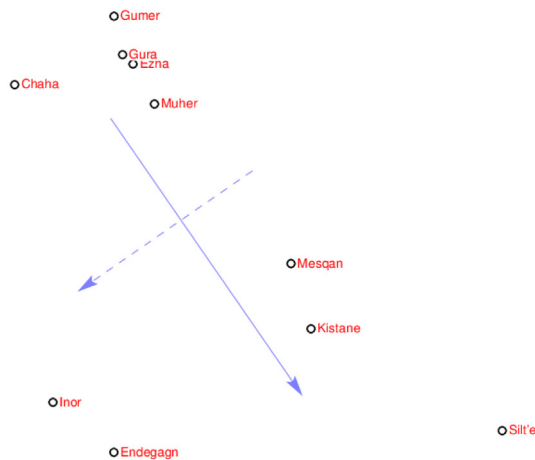
Fig. 5. Classification of South Ethiopese languages based on structural distances.

{Mesqan and Muher} form another group. {Kistane} and {Silt'e} are separate languages. Fig. 5(d) presents the dialect map of the language varieties, based on lexical distance. Unlike phonetic distance, there are five distinct groups of languages. Clearly, the phonetic and lexical classifications are different. For example, Endegagn and Inor form a group in the lexical classification, but not in the phonetic classification. Kistane and Silt'e are different in both classifications.

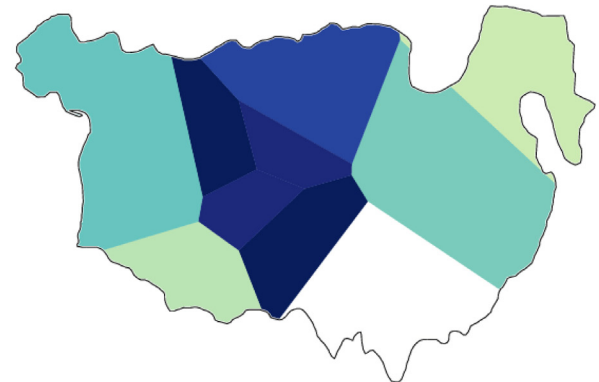
4.1.2. Classifications based on functional distance

The functional distance results were obtained from the Word Categorization test. Since the Word Categorization test measures the similarity, not the difference, among the language varieties, the average of the participants' scores on the test was subtracted from 100 to obtain functional distance. Fig. 6(a) presents a plot of multidimensional scaling of functional distance in two-dimensional spaces. The first dimension (solid arrow) shows that Silt'e has the highest value whereas Gumer, Chaha, Ezha and Gura have the lowest values. Muher and Mesqan have medium values. This dimension explains 79% ($r = .89$) of the variance in the distance matrix. The second dimension shows that Inor and Endegagn have the highest values, while Muher and Mesqan have the lowest values. This dimension explains 14% ($r = .37$) of the variance in the distance matrix. The two dimensions together explain 93% of the variance in the distance matrix. Fig. 6(b) illustrates the map of the first dimension of the multidimensional scaling. The lightest color indicates area

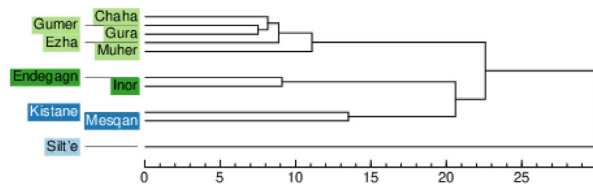
A PREPRINT - APRIL 22, 2020



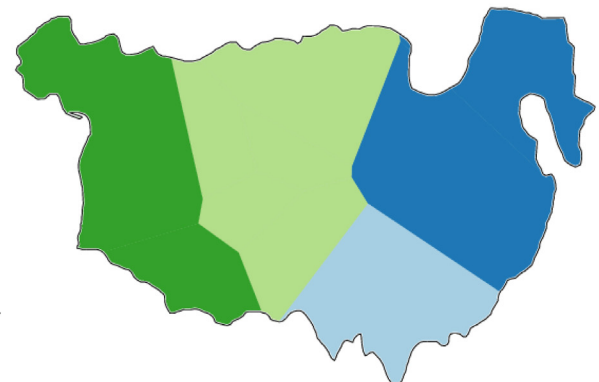
(a) Plot of multidimensional scaling for the functional distance, D1 = 79%, D2 = 14%



(b) Multidimensional scaling map for the functional distance. The light color shows area with the highest distance value



(c) Classification of South Ethiiosemitic languages based on functional distance



(d) Dialect map of South Ethiiosemitic languages based on functional distance

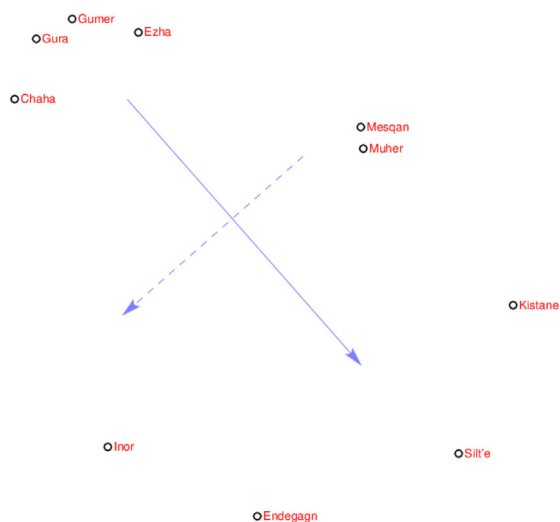
Fig. 6. Classification of Ethiiosemitic languages based on functional distances.

with the highest distance value which is the Silt'e area. The pattern in the multidimensional scaling plot shows that there are roughly five groups of the language varieties – {Gumer, Chaha, Ezha and Gura} form one group, {Muher, and Mesqan} another group, and {Inor and Endegagn} also form another group. {Silt'e} and {Kistane} are separate languages. Fig. 6(c) and 6(d) present the dendrograms of the language varieties based on functional distance, and the corresponding dialect map. As can be seen from Fig. 6(c), {Gumer, Gura, Chaha and Ezha} form a group. {Muher and Mesqan} form another group. Moreover, {Endegagn, Inor} are closely related. {Silt'e} and {Kistane} are separate languages. Fig. 6(d) also shows five language areas, with Silt'e and Kistane forming their own distinct dialect area.

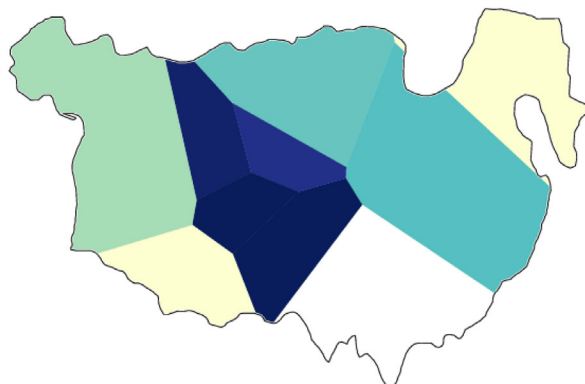
4.1.3. Classifications based on perceptual distance

In Section 3 it was indicated that two perceptual distance measures, that is, perceived similarity and perceived intelligibility, were employed to determine perceptual distance among the language varieties. The percentage of the mean of the two measures was computed and subtracted from 100 to quantify perceptual distance among the varieties. It is important to remember that the perceptual test measures the similarity among the language varieties, not the difference, and this is why the subtraction was needed. The cluster analysis was performed on the average of the upper and the lower halves of the perceptual distance matrix.

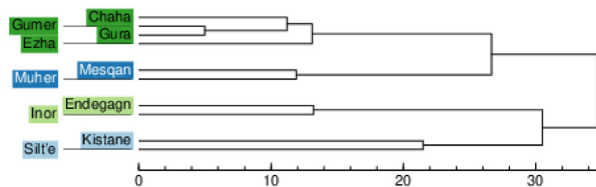
Fig. 7(a) shows the multidimensional scaling plot of perceptual distance. As the figure illustrates, in the first dimension, Ezha, Gumer, Gura and Chaha have the lowest values while Kistane and Silt'e have the highest values. This dimension



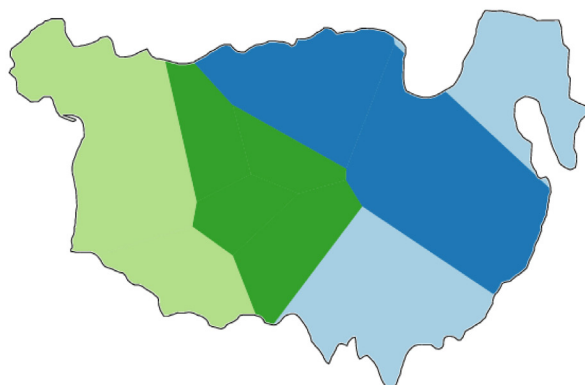
(a) plot of multidimensional scaling for the perceptual distance, D1 = 76%, D2 = 7%



(b) Map of the first dimension of multidimensional scaling in two dimensional space, for the perceptual distance. The light color shows the area with the highest distance value.



(c) Classification of South Ethiiosemitic languages based on perceptual distance



(d) Dialect map of Ethiiosemitic languages based on perceptual distance

Fig. 7. Classification of South Ethiiosemitic languages based on perceptual distances.

explains 76% ($r = .87$) of the variance in the distance matrix. The second dimension (dashed arrow) shows that Inor has the highest value while Mesqan and Muher have the lowest values. This dimension explains 7% ($r = .27$) of the variance. The remaining values are between these two extremes. Both dimensions combined explain 83% of the variance in the distance matrix. The multidimensional scaling results clearly show that there are four groups of language varieties: {Chaha, Gura, Gumer and Ezha}, {Mesqan and Muher}, {Endegagn and Inor} and {Kistane and Silt'e}. From a perceptual point of view, Kistane is closely related to Silt'e. Fig. 7(b) show the map of the first dimension of the multidimensional scaling. The light color shows the area that has the highest linguistic distance which is the Silt'e area. Fig. 7(c) and (d) show the classification of the languages based on perceptual distance, and the dialect map of the 10 language varieties respectively. Fig. 7(c) shows that {Chaha, Gumer, Gura and Ezha} form a group. {Inor and Endegagn} form a group. There is also a strong affinity between Muher and Mesqan. In a different manner from the classifications based on structural and functional distances, Kistane and Silt'e form a group in the classification based on perceptual distance. Fig. 7(d) shows the dialect map of the South Ethiosemitic languages based on the perceptual measure.

4.1.4. The combined classification of Ethiosemitic languages

As presented in the preceding sections, the classifications that were obtained from the structural, functional and perceptual distance measures are not identical. The classification based on phonetic distance shows six groups of languages, while the classification based on lexical distance indicates five groups. Hence, this section aims to combine these classifications and provide a comprehensive classification of the languages. The results of the comparison between the combined classification and the classifications by the historical linguists will then be presented. Fig. 8(a)-(d) summarizes the classifications presented in table 4.1.1 to 4.1.3. Fig. 8(e) presents the combined classification which was derived from the comparisons of all other classifications. The Sigma symbol in the combined classification represents an unspecified mother language.

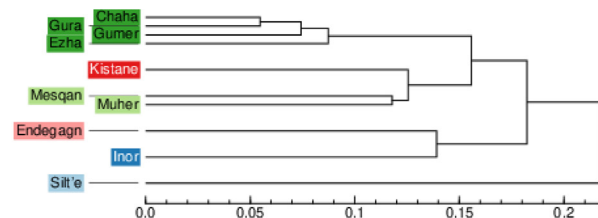
Given that the linguistic distance was measured from three perspectives (structural, functional and perceptual), the distance matrices were ranked based on their reliability, and the most reliable distance measures were prioritized in the process of combining the classifications presented above. Gabmap provides two measures of reliability of distance matrices: local incoherence and Cronbach's alpha. Local incoherence is a numerical score of local stress that is assigned to a set of differences between items (a measure of linguistic distances in the present study). The optimal score is zero while the non-optimal scores are any positive value. Comparing the value of local incoherence for different measurements over the same data indicates which result is more reliable (Nerbonne et al., 2011). Lower values of local incoherence mean that the results are better. The idea behind local incoherence is that, on average, the locations that are close should be less different than locations that are further apart.

Cronbach's alpha is a coefficient of reliability. It is used to measure the internal consistency or reliability of the psychometric test scores. In Gabmap, it is used as the coefficient of reliability of the measurement of differences over the data. High ($>.70$) Cronbach's alpha values mean that there is a high level of consistency among the measure of distances. Table 2 shows the results of local incoherence and Cronbach's alpha for each of the distance matrices: phonetic, lexical, functional and perceptual.

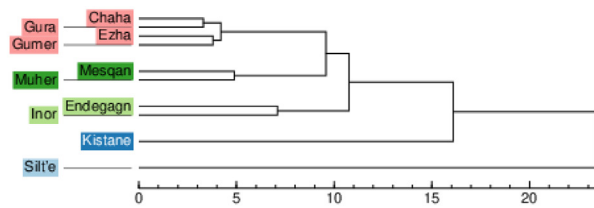
Table 2 shows that phonetic distance has the highest Cronbach's alpha value, and the lowest value of local incoherence. This means that it is the most reliable measure compared to all other distance measures. Lexical distance has lower local incoherence and a higher Cronbach's alpha compared to the functional and perceptual distance measures. Compared to perceptual distance, functional distance has a high Cronbach's alpha and a low value of the local incoherence. Perceptual distance has the lowest Cronbach's alpha and the highest local incoherence, which means that it has very low reliability. In general, Table 2 shows that structural distance (both phonetic and lexical measures) presents the most reliable distance measures. Functional distance is more reliable than perceptual distance. Perceptual distance is the least reliable distance measure.

Given these differences in reliability, structural distance was employed as a primary parameter in the process of determining the combined classification, that is, if a set of language varieties form a group in both phonetic and lexical classifications, that set of languages was automatically considered for the combined classification. However, when languages belong to different groups in the phonetic and in the lexical classification, functional distance was considered as a second parameter to determine which group is the most plausible one. Perceptual distance was considered as a third parameter when a set of language varieties formed different groups in the classifications based on both the structural and functional distances.

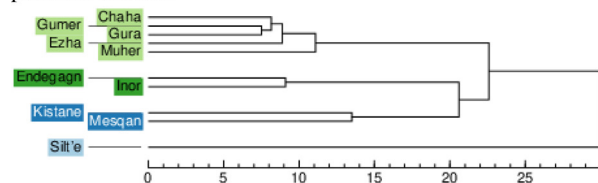
In Fig. 8(a)-(d), {Chaha, Gura, Ezha and Gumer} form a group not only in the classification based on phonetic distance, but also in the classification based on lexical distance. Therefore, this group was automatically included in the combined classification without even considering their classification based on functional and perceptual measures. {Inor} and {Endegagn} are separate languages in the classification based on phonetic distance, but they are very similar in the classification based on lexical distance. Therefore, functional distance was used as a second parameter. Based on these



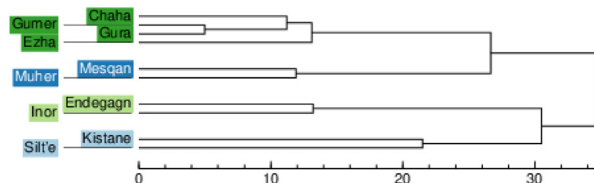
(a) Classification of South Ethiosemitic languages based on phonetic distance



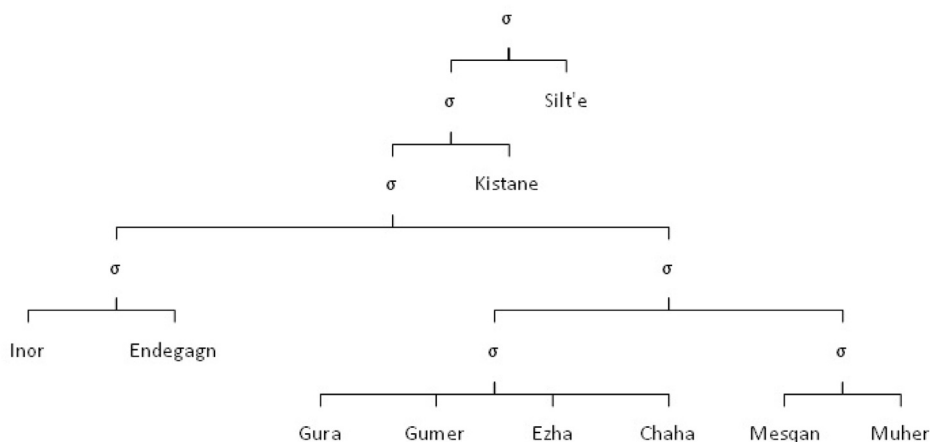
(b) Classification of South Ethiosemitic languages based on lexical distance



(c) Classification of South Ethiosemitic languages based on functional distance



(d) Classification of South Ethiosemitic languages based on perceptual distance



(e) Combined classification of South Ethiosemitic languages

Fig. 8. Comparisons of the classifications of South Ethiosemitic languages.

Table 2
Consistency within the distance matrices.

		Local incoherence	^a Cronbach's alpha
Structural	Phonetic	.22	.97
	Lexical	.23	.87
Functional		.29	.63
Perceptual		.32	.61

^a The high Cronbach's alpha of the phonetic distance could be due to the high sample size. Nonetheless, the higher degree of Cronbach's alpha of the remaining two measures (lexical and functional) clearly shows that perceptual distance has low reliability. It is also important to remember that the reliability measures for the functional and perceptual distances is based on the mean of the upper and the lower halves of the respective distance matrix.

criteria, Inor and Endegagn were grouped together in the combined classification. {Mesqan and Muher} form a group in the classifications based on both phonetic and lexical measures. Hence, they automatically qualified for the combined classification. {Silt'e} and {Kistane} are separate languages in the classification based on the phonetic and lexical parameters. They are also separate languages in the classification based on functional distance. Therefore, they were considered as independent languages in the combined classification although they form a group in the classification based on perceptual distance. This was due to the fact that perceptual distance has very low reliability. Based on these requirements, the 10 South Ethiiosemitic language varieties were classified into five groups – the first group consists of {Chaha, Gura, Gumer, Ezha}; the second group contains {Inor, Endegagn}, the third group comprises of {Mesqan, Muher}; the fourth group includes only {Kistane}, the fifth group consists of {Silt'e}.

As can be seen from Fig. 8(a)–(c), the grouping of the four Central West Gurage languages – Chaha, Gura, Gumer and Ezha is consistent across all the classification parameters. Therefore, the four Central West Gurage languages were used as a point of reference to determine the relative positions of other groups of languages in the combined classification. {Muher and Mesqan} are closer to {Chaha, Gura, Gumer and Ezha} than {Kistane} in the classification based on lexical distances. This is not the case for the classification based on phonetic distance since {Kistane} is rather close to {Chaha, Gura, Gumer and Ezha}. In this case, functional distance cannot be used as a second parameter since Muher and Mesqan do not form a group in the classification based on functional distance. Hence, perceptual distance was used as a third parameter to move {Muher and Mesqan} closer to the four Central West Gurage languages. {Inor, Endegagn} are closer to the Central West Gurage languages than {Kistane} in lexical, functional and perceptual classifications; therefore, they maintained their position in the combined classification. Moreover, compared to Silt'e, {Kistane} is closer to the Central West Gurage languages based on phonetic, lexical, functional and perceptual parameters. Silt'e is most remote from the Central West Gurage languages based on three (lexical, functional and perceptual) of the four classification parameters. The ultimate result of this process is the combined classification presented in Fig. 8(e).

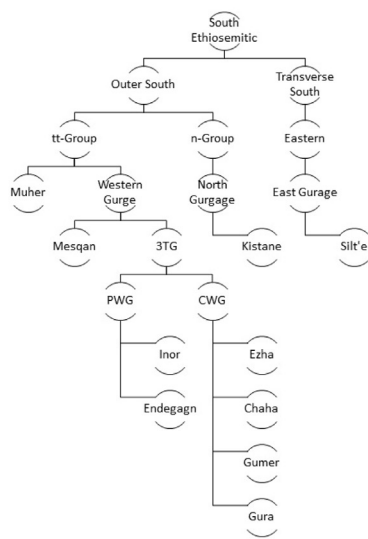
The remaining point is determining to what extent the combined classification corresponds to the classifications previously proposed by historical linguists. Fig. 9(a)–(c) shows that the combined classification seems similar to the classification by Hetzron (1972). For example, in both classifications, Chaha, Gura, Gumer and Ezha form a group. Inor and Endegagn also form a group in both classifications. However, unlike the combined classification, Muher and Mesqan do not form a group in the classification by Hetzron (1972). Moreover, unlike the classification by Demeke (2001), Muher and Inor do not form a group with the Central West Gurage languages which are {Chaha, Gura, Gumer and Ezha} in the combined classification.

Mere impressionistic comparisons of the dendrograms, may not precisely convey to what extent these classifications are similar. Hence, it is important for our purposes that the clustering procedures result in a re-estimation of the distances between collection sites, the so-called cophenetic distance (Nerbonne, 2010, p. 483). The cophenetic distance is the distance between two sites at the point at which they are fused in the clustering process. Cophenetic distances distort the original distance matrix because of the stipulation that the distance between the newly fused nodes, and all others, be the average of the distances from each component of the fusion to the others. For example, in Fig. 9(c), the cophenetic distance between Muher and Mesqan is two: (1) from Muher one node up to the mother node, (2) from the mother node down to Mesqan (see also Gooskens and Heuven, 2018). Pearson's correlation coefficient was used to illustrate the relationship between the cophenetic distance of the combined classification presented in Fig. 9(e) and that of the classifications by the historical linguists.

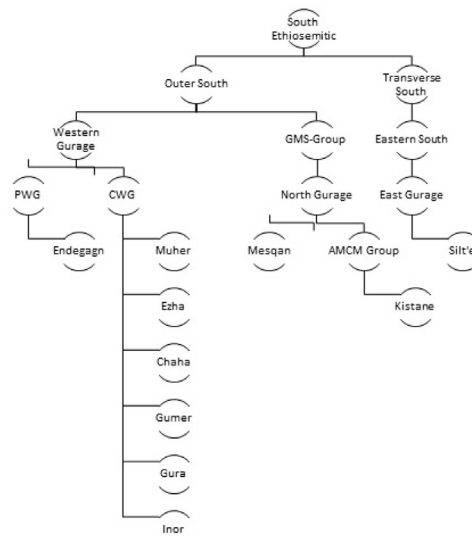
For the sake of simplicity and space, only the 10 language varieties under investigation are included in Fig. 9 among several Ethiiosemitic languages previously classified by the historical linguists. Since the distance between the nodes in a family tree is symmetrical (the distance between node A and node B is equal to the distance between node B and node A), the number of pairs of cophenetic distance measures is always $N \times (N - 1)/2$. This means that in the present study, there are 10 language varieties. Therefore, the possible symmetric pairs of languages to which the cophenetic distance has to be computed is $10 \times (10 - 1)/2$, which is 45. For the sake of space, only the correlation coefficients between the cophenetic distance of the combined classification and that of the classifications by Demeke (2001) and Hetzron (1972) are presented here. The analyses of the relationship using Pearson's correlation show that the cophenetic distance of the combined classification correlates more strongly to the cophenetic distance of the classification by Hetzron (1972), $r = .761$ as compared to the correlation between the cophenetic distance of the combined classification and that of the classification by Demeke (2001), $r = .553$. The two correlation coefficients are statistically significantly different, Hotelling's t -test, $t = 6.845$, $p < .001$.

4.2. Relations among the three dimensions of distance

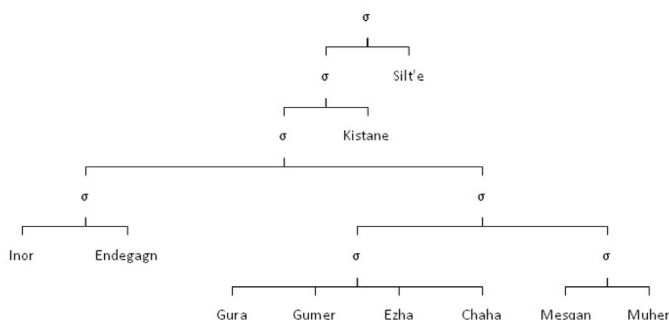
As indicated in Section 1.1, examining the relationship among the three dimensions of linguistic distance is one of the aims of the present study. In this section, therefore, correlations among the three dimensions of linguistic distances



(a) Classifications of the South Ethiiosemitic language according to [Hetzron \(1972\)](#)



(b) Classifications of the South Ethiiosemitic languages according to [Demeke \(2001\)](#)



(c) The combined classifications of the South Ethiiosemitic languages

Fig. 9. Comparisons between the combined classification and the classifications by historical linguists.

reported in the preceding sections are presented. [Table 3](#) illustrates the correlation coefficients of the two structural distances, functional distance and perceptual distance. As can be seen from the table there is a very strong correlation between the two structural distances – phonetic distance and lexical distance. Furthermore, the correlation between the two structural distances and perceptual distance is very strong. Compared to other correlation coefficients, the correlation between functional distance and perceptual distance is small (although not statistically significant). This suggests that the participants' similarity judgment and their actual score on the intelligibility test may not be exactly the same. In general, there are strong correlations among almost all the distance measures compared in [Table 3](#). As a result, in [Table 4](#), these correlation coefficients are compared to each other to determine whether there are statistically significant differences among them.

Fisher's r to z transformation was employed to compare the correlation coefficients among the three distance measures: structural, functional and perceptual. [Table 4](#) illustrates that there are no statistically significant differences among the correlation coefficients of all the distance measures.

Table 3
Correlation coefficients of the three dimensions of distance.

	Structural		Functional ^a	Perceptual
	Phonetic	Lexical		
Structural	Phonetic	.874	.804	.853
	Lexical		.849	.777
Functional				.747

Table 4
Comparison of the correlation coefficients.

Compared coefficients ^a		
	z-values	p.value
$r_{PcpD}r_{PD}$ VS. $r_{PcpD}r_{LD}$	1.051	.293
$r_{FD}r_{PD}$ VS. $r_{FD}r_{LD}$	-.654	.513

^a PcpD = perceptual distance, LD = lexical distance, FD = functional distance, PD = phonetic distance.

4.3. Intelligibility among the South Ethiosemitic languages

As indicated in Section 1, both functional distance and the degree of degree of intelligibility to be discussed in this section refer to the respondents' scores on the Word Categorization test. In other words, the respondents' scores on the Word Categorization test were used as a tool to indicate the degree of functional distance among the 10 South Ethiosemitic language varieties as well as to determine the degree of intelligibility among the language varieties. In this section, the respondents' scores on the Word Categorization test are presented. In Section 1, the intelligibility was defined as the degree of communication or understanding between the speakers of related languages, in principle, without having had a direct exposure to either of the languages. The assumption in the present study was that the correct categorization of the words into their semantic categories measures the degree of understanding (at least at the lexical level) of the speakers of the language varieties.

To determine the degree of intelligibility among the language varieties, a 75% intelligibility threshold was set based on the suggestion of Grimes (1995, p. 22) and partly based on the conservative nature of the test administered. Hence, a 75% or more score in the Word Categorization test was considered as confirmation of intelligibility between the test language and the language of the test-takers. 71–74% score was considered as partial intelligibility. Anything below 71% was considered as absence of intelligibility. Table 5 shows the intelligibility scores of the participants on the Word Categorization test.

As can be seen from Table 5, Chaha speakers understand Ezha, Gumer and Gura. Endegagn speakers partially understand Inor. Speakers of Ezha understand Chaha and Gumer. In the same manner, Gumer speakers understand Chaha, Gura, Ezha and Muher. Gura speakers understand Chaha, Ezha, Gumer and Muher. Inor speakers partially understand Chaha and fully understand Endegagn. Furthermore, Mesqan is partially intelligible to Ezha. Muher speakers understand Chaha. Silt'e and Kistane are not intelligible to any of the language varieties.

Table 5 further shows that the test-takers did not score 100% on their own native languages although, in principle, it is assumed that the native speakers have a perfect knowledge of their own language. The participants underperformed on their native languages probably due to non-linguistic factors such as fatigue, the quality of the recordings, lack of attention, noises in the test environment and time pressure. In order to compensate for the influence of these factors, adjusted means were computed for the participants' scores on the Word Categorization test. It was computed by subtracting the actual mean of the participants' score on their own native language from the hypothetical mean, which is always 100%. The mean differences were then added to the same participants' scores on the non-native languages with the assumption that the factors that affect the participants' scores on their native languages equally affect their scores on the non-native languages. For instance, Chaha speakers, on average, scored 81% on their own native languages, although they are supposed to score 100%. Therefore, the adjusted mean was computed by subtracting 81% from 100%, which is 19%.

Table 5

Mean of the participants' scores on the Word Categorization test.

Language ^a	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	81 ^b	58	81	85	81	69	50	46	69	42
Endegagn	62	81	48	48	43	71	48	43	57	33
Ezha	80	52	80	76	76	52	36	40	76	40
Gumer	82	54	79	86	82	50	57	68	82	36
Gura	83	52	79	83	86	55	59	59	79	38
Inor	71	91	64	68	55	82	50	45	55	32
Kistane	48	48	39	57	48	39	83	52	35	22
Mesqan	67	42	71	67	42	42	67	85	63	33
Muher	77	38	69	69	65	46	65	42	81	23
Silt'e	43	43	48	57	43	22	35	35	48	87

^a The test languages are abbreviated - CH = Chaha, ED = Endegagn, EZ = Ezha, GM = Gurmer, GU = Gura, IN = Inor, MS = Mesqan, MU = Muher, SI = Silt'e and KS = Kistane; the intelligibility results are converted to percentage.

^b The participants did not fully understand their own variety. This could be because of various factors including recording quality, time pressure, and lack of attention.

Then 19% was added to the scores of the Chaha participants on all other language varieties. Table 6 presents the adjusted mean scores computed based on the results illustrated in Table 5.

Based on the adjusted means presented in Table 6, Chaha speakers can understand Endegagn, Ezha, Gumer, Gura, Inor and Muher. Endegagn speakers can freely communicate with Chaha, Inor and Muher speakers. Speakers of Ezha understand Chaha, Gumer, Gura and Muher. They also partially understand Endegagn and Inor. Gumer speakers understand Chaha, Ezha, Gura, Mesqan and Muher. They also partially understand Kistane. Gura speakers understand Chaha, Ezha, Gumer and Muher. They also partially understand Kistane and Mesqan. Inor speakers understand Chaha, Endegagn, Ezha, and Gumer. They also partially understand Gura and Muher. Moreover, Mesqan speakers understand Chaha, Ezha, Gumer, Kistane and Muher. Muher speakers understand Chaha, Ezha, Gumer and Gura. Silt'e is not intelligible to any of the language varieties.

Menuta (2015, p. 189) argues that the best center of communication is Mesqan, based on the study he conducted on six Gurage varieties – Chaha, Inor, Kistane, Mesqan, Muher and Wolane. In other words, according to Menuta (2015), many speakers of Gurage varieties understand Mesqan better than the remaining Gurage varieties investigated in the study. The present finding contradicts this report. As can be seen in Fig. 10, it is Chaha that seems to be the center of communication. Chaha is intelligible to seven of the 10 language varieties investigated. Silt'e was excluded from the figure since it was not intelligible to any of the language varieties. In Fig. 10, the two-directional arrow shows that intelligibility is symmetrical, while the one-directional arrow shows that intelligibility is asymmetrical.

The difference between these two findings might be due to various factors. First, the present study only used the Semantic Word Categorization test. The authors recognize that testing intelligibility at a higher linguistic level may yield

Table 6

The adjusted mean of the test-takers' scores on the Word Categorization test.

Language ^a	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	100	77	100	100	100	88	69	65	88	61
Endegagn	81	100	67	67	62	90	67	62	76	52
Ezha	100	72	100	87	96	72	56	60	96	60
Gumer	96	68	93	100	96	64	71	82	96	50
Gura	97	66	93	97	100	69	73	73	93	52
Inor	89	100	82	86	73	100	68	63	73	50
Kistane	65	65	56	74	65	56	100	69	52	39
Mesqan	82	57	86	82	57	57	82	100	78	48
Muher	96	57	88	88	84	65	84	61	100	42
Silt'e	56	56	61	70	56	35	48	48	61	100

^a The test languages are abbreviated – CH = Chaha, ED = Endegagn, EZ = Ezha, GM = Gurmer, GU = Gura, IN = Inor, MS = Mesqan, MU = Muher, SI = Silt'e and KS = Kistane; the results are converted to percentage.

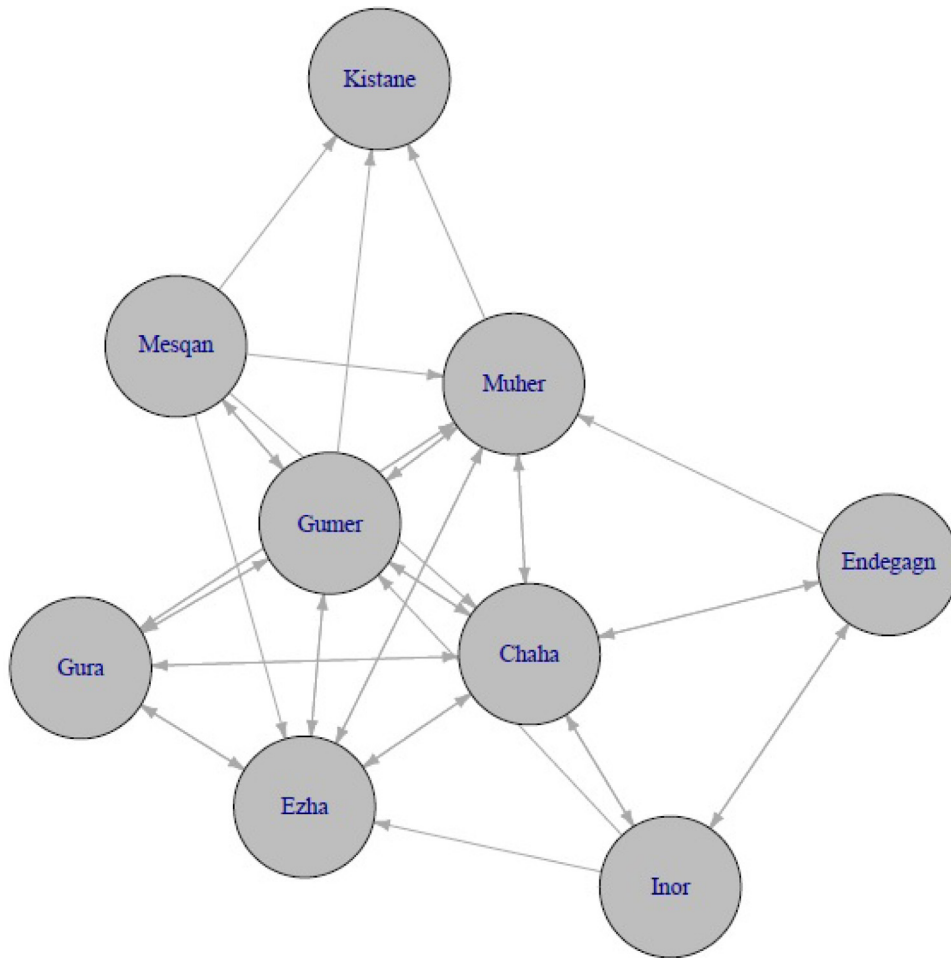


Fig. 10. Chaha is selected as the center of communication.

different results. Nonetheless, the present study opted for the inclusion of a relatively large number of languages and examine them from different perspectives, rather than just focusing on intelligibility. In this regard, [Menuta \(2015\)](#) included several tests which are promising. Nonetheless, there are also concerns about the approaches taken by [Menuta \(2015\)](#). We suspect that the priming effect was not properly controlled for; the same test materials were repeated across the speakers of different varieties; therefore, it is possible that the intelligibility scores were inflated because of the participants' familiarity with the test materials. Besides, [Menuta \(2015\)](#) tested elderly people while the participants of the present study were secondary school students. It could be that elderly people performed better on some of non-native languages as compared with the young people mainly because of their lifelong exposure to the non-native language varieties. Sample size could also be another factor. [Menuta \(2015\)](#) tested 12 participants from each site. The present study tested 30 participants from each site. A carefully selected small sample size could reflect exceptional performance of the participants because of their linguistic abilities. Moreover, during test administration [Menuta \(2015\)](#) asked the participants to provide written answers. It is not clear how the respondents managed to provide written answers since none of the Gurage varieties (except Silt'e) has a writing system.

5. Discussion

As presented in Section 4.2, the comparisons among the measures of the three dimensions of distance show that the two structural distances (phonetic and lexical) strongly correlate with each other. This implies that the two structural measures can be used interchangeably to determine the linguistic distance among related languages. The present study

also reported very strong correlation between structural distance and functional distance, although different materials were used to measure the two dimensions of distance. This suggests a high degree of substitutability between the two dimensions of measuring linguistic distance. Moreover, the strong correlation between structural distance and functional distance indicates that the respondents' scores on the intelligibility test have a strong connection with the properties of the structure of the language varieties.

Given that there is no significant difference between the correlation coefficient of phonetic distance and intelligibility scores, and that of lexical distance and intelligibility scores, it seems that there is no difference between the two structural distances in terms of their influence on the participants' scores on the Word Categorization test. This finding is slightly different from previous studies, which reported a stronger correlation between lexical distance and functional distance as compared to the correlation between phonetic distance and functional distance (e.g., [Tang and van Heuven, 2009](#)). It also differs from studies that reported a stronger correlation between phonetic distance and functional distance, but not between lexical distance and functional distance (e.g., [Van Bezooijen and Gooskens, 2007](#)). Many factors, such as similarity of phoneme inventory and word frequency, may contribute to the relationship between functional distance and structural distance. The relationship between these two dimensions is probably language specific. For instance, in some languages, lexical similarity might be more important than phonetic similarity, while in other languages a slight phonetic difference may lead to misunderstanding. Moreover, the strong correlation between structural distance and perceptual distance shows that perceptual distance can be used as an alternative means of determining the linguistic distance among related languages, especially in a situation where gathering real linguistic data is difficult. Similar results were previously reported by [Gooskens and Heeringa \(2004\)](#) and by [Tang and van Heuven \(2009\)](#). This is good news for less studied languages that do not have dictionaries or detailed descriptions of their linguistic features. However, the low level of consistency in the perceptual distance matrix hints that there is a risk of using a mere perceptual distance to measure linguistic distance among related languages. This is because the perceptual perspective of measuring linguistic distance is more subjective than other means of measuring linguistic distance. As noticed by [Golubović and Sokolić \(2013\)](#), [Abu-Rabia \(1996\)](#), [Abu-Rabia \(1998\)](#) and [Pavlenko \(2006\)](#), the impact of language attitude is also more pronounced in situations where there are political divisions, stereotyping, and social and cultural hostilities.

Furthermore, the close similarity between the classifications based on the three dimensions of distance and the genealogical classifications previously provided by the historical linguists implies that, in addition to structural distances, functional and perceptual distances can also be used to classify related languages. In the present study, we noticed very close similarity between typological classifications and genealogical classifications. This result is consistent with a previous report by [Tang and van Heuven \(2009\)](#). In general, the correlations among the three dimensions of distance, which are reported in the present study, are consistent with studies previously conducted on Scandinavian languages (e.g., [Gooskens, 2005](#); [Gooskens, 2007](#); [Gooskens and Heuven, 2018](#); [Gooskens and Heeringa, 2004](#)) and on Chinese dialects ([Tang and van Heuven, 2009](#); [Tang and van Heuven, 2009](#)). These studies, in general, indicate that the distance among related languages can be measured from different perspectives. It is up to a researcher to choose the right perspective based on various factors, such as resources at disposal and the desired study objectives; for example, whether the aim of the study is typological or genealogical classification. Our study partly supports the claim that non-linguists' level of awareness can be used as a valid means of measuring distances among related languages, but we are also cognizant of the enduring debate regarding the validity of the perception-based approach (see [Goeman, 1999](#)) because of the low reliability of the perceptual distances we observed.

The classifications of Ethiosemitic languages based on the results obtained from the structural, functional and perceptual distance measures show that Chaha, Ezha, Gumer and Gura are very closely related languages. Mesqan and Muher also have very strong lexical affinity with these four languages. The lexical affinity among these language varieties was also reported in [Menuta \(2015\)](#). Mesqan and Muher have also close phonetic and lexical similarities. Kistane and Silt'e are different from all the remaining language varieties. This difference is probably due to the influence of the Cushitic languages on Silt'e and Kistane. This is an intuitive suggestion, and the interaction between South Ethiosemitic languages and the surrounding Cushitic languages is an issue that future studies may address.

The comparisons of the classifications obtained from the three distance measures show that the South Ethiosemitic languages under investigation can be classified into five groups. {Chaha, Gura, Gumer and Ezha} form a group. {Muher and Mesqan} are very similar languages; hence, they form a second group. {Inor and Endegagn} consistently form the third group. {Kistane} and {Silt'e} are different from all other language varieties. These classifications are very similar to classifications previously proposed by [Hetzron \(1972\)](#), but somehow differ from others, for example, [Demeke \(2001\)](#). [Demeke \(2001\)](#) classified Mesqan under North Gurage, together with Kistane. Although both the structural and functional measures show that Kistane and Silt'e are different languages, the speakers of the language varieties believe that their languages are similar to each other. The causes of the mismatch between the speakers' perceptions and the linguistic reality merit further investigation.

With regard to the intelligibility among South Ethiosemitic languages, the results obtained from the Word Categorization test show that Chaha, Gura, Gumer and Ezha are mutually intelligible. Muher and Mesqan are partially

intelligible with these languages. This partial intelligibility is slightly different from the full intelligibility previously reported in Menuta (2015). Endegagn and Inor are also mutually intelligible. Kistane and Silt'e are not intelligible with any of the Gurage varieties investigated in the present study. The reported intelligibility scores are largely asymmetrical. As noticed by Gooskens (2018), Gooskens et al. (2010) and Gooskens (2007), this asymmetry might be due to linguistic and non-linguistic factors. Some languages can be incomprehensible because of their complicated phonological structures, such as pervasive reductions due to assimilation, and alternation between obstruents and approximants (Bleses et al., 2008, p.623). Gooskens (2018) has also discussed various non-linguistic variables, such as contact and experience, orthography, gesture and language attitude.

Based on the findings presented in Section 4, we also provide our position as to whether the South Ethiosemitic languages investigated in the present study are dialects or not. While providing a clear-cut boundary between 'dialect' and 'language' is always difficult due to various linguistic and non-linguistic factors, the results of the cluster analyses and the intelligibility scores suggest that Silt'e and Kistane are different languages. The remaining languages are dialects of the same language. Determining whether these varieties are dialects or independent languages may have significant consequences for the attempts that have been made to standardize the language varieties. The results of our study imply that Kistane and Silt'e need to be treated as separate languages in the standardization process. The remaining Gurage varieties can be considered as dialects, and the same materials can be used to employ these languages for schooling, media and different administrative purposes.

Appendix A. Data collection tools

A.1. Response sheet for words categorization test

Instruction: Dear student, you are going to listen to some list of words. Listen carefully and determine in which of the following categories each word belongs. For one word there is only one possible category. Provide your answer by putting 'X' mark in the box provided in front of each category. Note that for every audio stimulus, there are 10 options of word categories.

- | | |
|----------------------|--------------------------|
| 1. Cloths | <input type="checkbox"/> |
| 2. Body Parts | <input type="checkbox"/> |
| 3. Kitchen Utilities | <input type="checkbox"/> |
| 4. Fruits | <input type="checkbox"/> |
| 5. Food Type | <input type="checkbox"/> |
| 6. Domestic Animals | <input type="checkbox"/> |
| 7. Furniture | <input type="checkbox"/> |
| 8. Vegetables | <input type="checkbox"/> |
| 9. Wild Animals | <input type="checkbox"/> |
| 10 .Cereals | <input type="checkbox"/> |

Appendix B. Test materials

B.1. List of words used for determining phonetic and lexical distance

No	English	Amharic	Oromo	Chaha	Endegagn	Ezha	Gumer	Gura	Inor	Kistane	Meskan	Muher	Silt'e
1	added	dammara ¹	dabale ¹	daparom ¹	dappara ¹	dabbaram ¹	dapparam ¹	daparom ¹	dappara ¹	dabbalam ¹	dabbaram ¹	dabbaram ¹	dabala ¹
2	all	hullum ¹	hunda ¹	innim ¹	hin?edi ²	innim ¹	innim ¹	innim ¹	inna ¹	kullam ¹	innim ¹	innim ¹	hullimko ¹
3	ape	fofa ¹	kamale ²	wanka ¹	wenka ¹	wank'a ¹	wank'a ¹	wanka ¹	wanke ¹	kamale ²	kamale ²	kamale ²	kamale ²
4	arm	kind ¹	irree ³	agat ²	agad ²	hinna ¹	xina ¹	hina ¹	hini?o ¹	kirro ³	hinna ¹	hitte ¹	kire ³
5	back	garba	duuba	g'isa ¹	g'isa ¹	ginza ¹	f'ar ²	g'isa ¹	gisiz'o ¹	ginza ¹	ginza ¹	ginza ¹	cin ³
6	barley	gabs ²	garbuu ²	axir ¹	axir ¹	axir ¹	axir ¹	axir ¹	axir ¹	gabs ²	ihir ¹	axi ¹	ixal ¹
7	basil	bassobila ¹	maslobbaa ¹	masobilal ¹	masobila ¹	masobilal ¹	bassobila ¹	bassobila ¹	masobilal ¹	masobila ¹	k'asonat ¹	bassobila ¹	bakar ²
8	bean	adang'arre ¹	asongoree ¹	adeng'alle ¹	adeng'alle ¹	adeng'alle ¹	adeng'are ¹	adang'aro ¹	¹	bolokke ²	adang'arre ¹	adengurro ¹	boloke ²
9	beautiful	konjo	bareeddu ²	markama ¹	markama ¹	markama ¹	markamma ¹	markama ¹	markama ¹	markama ¹	markama ¹	markama ¹	bareedo ²
10	bed	alga ¹	sire	alga ¹	anz ¹	alga ¹	alga ¹	arga ¹	arga ¹	alga ¹	alga ¹	alga ¹	dugmal ²
11	began	gammara ²	eegale	kanasam ¹	kanasa ¹	kanasam ¹	kanasam ¹	kanasa ¹	kanasa ¹	kanasa ¹	kanasa ¹	kanasa ¹	gammara ²
13	belt	kabatto ¹	kabattoo ¹	matagača ²	kaboto?o ¹	kabatto ¹	kabatto ¹	kabatto ¹	zabotot ¹	kabatto ¹	kabatto ¹	kabatto ¹	kabatto ¹
14	bosso	basso ¹	basso ¹	basso ¹	basso?o ¹	bassiwa ¹	bassiwa ¹	bassiwa ¹	bassiwa ¹	bassiwa ¹	bassiwa ¹	bassiwa ¹	basso ¹
15	big	tillik	guddaa	nikeya ¹	nu?i ¹	nikeya ¹	nikeya ¹	nikeya ¹	nu?iya ¹	malak ²	nikeya ¹	nikeya ¹	nikeya ¹
16	black	tukur ¹	gurraača	tukur ¹	gambona ²	gambona ²	tukur ¹	tukur ¹	gambona ²	tukur ¹	tukur ¹	tukur ¹	tukur ¹
17	blew	naffasa ¹	afuufe	naffasom ¹	naffasa ¹	naffasom ¹	naffasom ¹	naffasom ¹	naffasa ¹	naffasa ¹	naffasa ¹	naffasom ¹	naffasa ¹
18	blood	dam ¹	duja	dam ¹	dam ¹	dam ¹	dam ¹	dam ¹	dam ¹	dam ¹	dam ¹	dam ¹	dam ¹
19	boiled	afalla ²	danfise	čakarom ¹	čakarra ¹	čokk'oro ¹	čokkorom ¹	čokoro ¹	čok'oro ¹	fallam ²	čokk'oro ¹	čokk'oro ¹	čokk'oro ¹
20	bone	at'im ¹	lafee	at'im ¹	at'im ¹	at'im ¹	at'im ¹	at'im ¹	at'im ¹	at'im ¹	at'im ¹	at'im ¹	at'im ¹
21	bread	dabbo ¹	daabbo ¹	dabbo ¹	dabo?o ¹	dabbo ¹	dabbo ¹	dabbo ¹	dabbo?o ¹	dabbo ¹	furno ²	dabbo ¹	furno ²
22	breakfast	kurs	čiree	ginzir ¹	ginzir ¹	ginzir ¹	ginzir ¹	ginzir ¹	ginzir ¹	ginzir ¹	yaddara ²	addara ²	ginzir ¹
23	breast	tut ¹	harma	tut ¹	tut ¹	tut ¹	tut ¹	tut ¹	tut ¹	tibuyyo ¹	tut ¹	tut ¹	tut ¹
24	brother	wandimn	obboleessa	g'appa ¹	əššə?əm ²	g'əbbe ¹	g'əppay ¹	g'əpa ¹	əsom ²	zəmmi ¹	g'əbbe ¹	g'əbbe ¹	indat ⁴
25	bull	bare ²	dibicča	wir ¹	wir ¹	wir ¹	wir ¹	wir ¹	wir ¹	wir ¹	wir ¹	wir ¹	wir ¹
26	cabbage	gommon	raafu	ambir ¹	ambir ¹	ambir ¹	ambir ¹	ambir ¹	ambir ¹	ambir ¹	ambir ¹	ambir ¹	ambir ¹
27	calf (n)	¹	gabbu	m'asa ¹	mosa ¹	m'assa ¹	m'asa ¹	m'asa ¹	m'asa ¹	t'əg ²	m'asa ¹	m'asa ¹	izək ³
28	came	mə'ta ²	dufe	čəna ¹	mə?a ²	čəna ¹	čəna ¹	čəna ¹	mə?a ²	mə'ta ²	čəna ¹	bəssam ³	mə'ta ²
29	came out	wə'tačč ¹	baate	wə'tam ¹	wə?ačč ¹	wə'taččim ¹	wə'taččim ¹	wə'ta ¹	wə?ačč ¹	wə'tatti ¹	wə'ta ¹	wə'taččim ¹	wə'ta ¹
30	camel	giməl ¹	gaala ¹	gamera ¹	giməl ¹	gamera ¹	gamera ¹	gamera ¹	gamera ¹	giməl ¹	gamela ¹	gamela ¹	gaameela ¹
31	cat	dammot	adurre ²	angača ¹	angačč ¹	angačč ¹	angačč ¹	angača ¹	angača ¹	angača ¹	angača ¹	angača ¹	aadon ¹
32	chair	wənbər ¹	barčuma ⁴	wənbər ¹	tə'kəša ²	wənbər ¹	wənbər ¹	wənbər ¹	tə'kəša ²	wənbər ¹	sorər ³	wənbər ¹	borčuma ⁴
33	chassis	rəkəbot ²	rəkəboot ²	yəsinačč ¹	rəkəbot ²	yəsinačč ¹	yəsinačč ¹	yəsinačč ¹	sine? ¹	rəkəbot ²	yəsinačč ¹	yəsinačč ¹	ganč ³
34	cheap	rikkaš ³	rakša ³	giwa ¹	wudo?o ²	giwa ¹	giwa ¹	giwa ¹	widi?o ²	arkus ³	rikkaš ³	giwa ¹	ruks ³
35	cheek	gunc ²	mallaa	danga ¹	danga ¹	danga ¹	danga ¹	danga ¹	gunc ²	gunc ²	danga ¹	gunc ²	gunc ²
36	chest	dərot ¹	lap'pe	data ¹	datto ¹	dadda ¹	datta ¹	fanika ²	data ¹	dərot ¹	hin ³	dadda ¹	wəzəna ³
37	chickpea	šimbura ¹	šumbura ¹	čəmb'ora ¹	šumbura ¹	čəmb'ora ¹	šimbura ¹	šimbura ¹	šimbura ¹	šəmbura ¹	šimbura ¹	šimbura ¹	šimbura ¹
38	child	hišan	daa?ima	tika ¹	tika ¹	tika ¹	tika ¹	tika ¹	tika ¹	bayy ²	tuxə ²	tika ¹	čilo ³
39	cloak(n)	kot ¹	koota ¹	kot ¹	kot ¹	kot ¹	kot ¹	kot ¹	kot ¹	kot ¹	kot ¹	kot ¹	kote ¹
40	cloud	damməna ¹	duumessa ¹	dabəra ¹	dawəna ¹	dabəra ¹	dabəra ¹	dabəra ¹	damara ¹	daməna ¹	dabəna ¹	dabəna ¹	dabəna ¹

41	coffee	bunna ²	buna ²	kawa ¹	kəwa ¹	buna ²	kawa ¹	kawa ¹	kawa ¹	buna ²	kawa ¹	kawa ¹	buno ²
42	coming	iyyməŋtʰa ¹	dufaa	yičən ²	yiməŋa ¹	tičən ²	yičən ²	tičən ²	timəŋa ¹	iyyməŋtʰa ¹	tišəŋi ²	tišəŋi ²	timəŋa ¹
43	(her self)	təbəxəččə ¹	dokotte	təšəməččə ¹	tədoŋəməččə ²	təšəməŋi ¹	təšəməčim ¹	təšəməččəŋi ²	tədoŋəməččəŋi ²	təsoŋəŋi ¹	təšəmə ¹	təšəməččim ¹	təšəmə ¹
44	cow	lam ¹	saʔa	aram ¹	anaw ²	aram ¹	aram ¹	aram ¹	aram ¹	anam ¹	anam ¹	anam ¹	lam ¹
45	cried	aləŋkəsə	booye	bəkəm ¹	bəkka ¹	bəkəm ¹	bəkkə ² am ¹	bəkə ¹	bəkə ¹	bəššəm ¹	bəkkə ¹	bəkkə ² am ¹	bəčə ¹
46	crocodile	azzo ¹	naaččə	zabəhəl ¹	dəwəwə ²	azzo ¹	azzo ¹	yəbərʔaw ¹	abarʔaw ¹	azzo ¹	azzo ¹	azzo ¹	wəwə ⁵
47	decide	məwəssan ¹	murteessu	wəwəssin ¹	iğgəwəŋi ²	fənnadəm ³	wəssot ¹	wəbərʔ	wəssin ¹	wəwəssin ¹	wəwəssin ¹	wəwəssin ¹	koočə ⁵
48	deer	agazən ¹	gadamsa	zomə ²	genzo ¹	gəzma ¹	sisəmə ⁴	imar ⁵	wəšir ³	fiččə ⁶	agazzən ¹	agazzən ¹	gūb ¹
49	destroy	mawdam	barbadeessu	banaram ¹	banarə ¹	banəssəm ¹	čifəčəfəm ²	banaram ¹	banarə ¹	dasəm ¹	barəssə ¹	afərrəsəm ⁴	dasəm ¹
50	did	adarrəgə	gofe	amənəm ¹	eppə ²	amənənəm ¹	amənəm ¹	epə ²	epə ²	kinnaw ³	šəkkəŋə ⁴	šəkkəŋə ⁴	anə ⁵
51	dispute(n)	fukukkir ³	falmii	šir ¹	manad ²	fukukkir ³	m ¹ agza ⁴	guwara ⁵	m ¹ agza ⁴	fukukkir ³	m ¹ agazza ⁴	fukukkir ³	fukukku ³
52	dog	wišša ²	sare	giyo ¹	giyo ¹	giyyo ¹	giyo ¹	giyo ¹	giyo ¹	wiŋša ²	giyo ¹	giyo ¹	buučo ³
53	donkey	ahiyya	harre	imar ¹	əwan ¹	imar ¹	imar ¹	imar ¹	imar ¹	əmar ¹	imar ¹	imar ¹	uumar ¹
54	door	bər ³	hulaa	wəfənča ¹	sanka ²	sanka ²	bər ³	wəfənča ¹	sanka ²	sanka ²	wəfənča ¹	wəfənča ¹	wəzəb ⁴
55	down	taččə ¹	gad ²	təŋaya ¹	təttinya ¹	wəšəŋiya ¹	wəšəŋiya ¹	gadaniya ²	təttinya ¹	təttinya ¹	wəttinya ¹	wəttinya ¹	uŋakolo ³
12	(v)(2.M.SG)	alləmə	manaabe	nəzəzəm ¹	nəzəzə ¹	nəzzəzəm ¹	nəzəzəm ¹	nəzəzəm ¹	nəzəzə ¹	birəzzəzə ¹	birəzzəzə ¹	nəzəzəm ¹	birəzzəzə ¹
56	dress (n)	kəmis ¹	uffata	kəmis ¹	kəmis ¹	muluʔiğ ²	kəmis ¹	kəmbis ¹	mardofa ³	dirbadir ⁴	kəmis ¹	kəmis ¹	kəmbissə ¹
57	drink (v)	fəŋfə	diğe	səčə ¹	seččə ¹	səččə ¹	siŋ ¹	səčə ¹	səčə ¹	səččə ¹	səččə ¹	səččə ¹	səččə ¹
58	dry (adj)	dərək ¹	gogaa	fərək ¹	dərə ²	fərək ¹	fərək ¹	fərək ¹	dərə ²	dərək ¹	fərək ¹	fərə ²	dərək ¹
59	ear	goro	gurra	inzi ¹	inzi ¹	inzi ¹	inzi ¹	inzi ¹	inzi ¹	inzi ¹	inzi ¹	inzi ¹	izni ¹
60	egg	inikulal ¹	bup p'a ²	inik'ira ¹	inikula ¹	anik'ə ¹	inik'ira ¹	inik'ira ¹	inik'ira ¹	aniko ¹	anik'ə ¹	anik'ə ¹	bup p'a ²
61	eight	simmint ¹	saddet ¹	sumt ¹	simut ¹	simut ¹	simut ¹	simut ¹	simut ¹	səmmint ¹	simmut ¹	simmut ¹	sumut ¹
62	elephant	zihon ¹	arba	ʔəxwə ¹	zəhon ¹	ʔəxwə ¹	ʔəxwə ¹	ʔəxwə ¹	ʔəxwə ¹	zəxon ¹	ʔəxwə ¹	zəxon ¹	dəxəno ¹
63	evening	miššir ²	galgala	gəpat ¹	gəppaʔad ¹	gəbbat ¹	gəppat ¹	gəbat ¹	gəpad ¹	mišə ²	gəbbat ¹	gəbbat ¹	gəbbat ¹
64	expensive	widd ²	kaali	fira ¹	mud ²	fira ¹	fira ¹	fira ¹	mumd ²	čəm ³	čum ³	fina ¹	čum ³
65	eye	ayn ¹	iğa	en ¹	en ¹	en ¹	en ¹	en ¹	er ¹	in ¹	en ¹	en ¹	in ¹
66	farmer	araas ²	koteebulaa	čəwəč ¹	čəwəč ¹	čəwəč ¹	čəwəč ¹	čəwəč ¹	čəwəg ¹	araš ²	araš ²	čəwəč ¹	arašš ²
67	father	abbat ¹	abbaa ¹	ab ¹	aba ¹	abba ¹	abba ¹	aβa ¹	aba ¹	abi ¹	ab ¹	ab ¹	abot ¹
68	fava bean	bakela ¹	baakilaa ¹	bakela ¹	bakela ¹	bakella ¹	bakella ¹	atarə ²	atarə ²	bakela ¹	bakela ¹	bakela ¹	bakela ¹
69	fenugreeek	abis ¹	hulbata	abəšə ¹	awəšə ¹	abəš ¹	abəšə ¹	abəšə ¹	abəšə ¹	abəšə ¹	abəšə ¹	abəš ¹	šuko ²
70	finger	fəŋfə	kuba	atebat ¹	aŋəya ¹	atebat ¹	atebat ¹	aŋəbə ¹	aŋəya ¹	aŋəbat ¹	aŋəbat ¹	aŋəbat ¹	in'abit ¹
71	finish(v)	čərrəssə ³	tumure	gəporə ¹	gəpporə ¹	fəgğom ²	gəpporə ¹	gəporə ¹	gəporə ¹	čərrəssəm ³	ʔəbbəro ¹	fəgğom ²	čərrəssə ³
72	five	amist ¹	šan	amist ¹	amist ¹	amist ¹	amist ¹	amist ¹	amist ¹	ammist ¹	ammist ¹	ammist ¹	ammist ¹
73	food-table	məsob ²	masooba ²	səf ¹	məsob ²	səfet ¹	məsob ²	səf ¹	səf ¹	lemat ³	lemat ³	wəskənbiya ⁴	iffə ¹
74	found	agaŋno ²	arge ²	nəkəbəm ¹	nəkka ¹	nəkəbəm ¹	nəkəbəm ¹	nəkəbə ¹	nəkəbə ¹	agaŋno ²	rəkəbə ¹	agaŋnəm ²	rəkəbə ¹
75	four	arat ¹	afur	arbat ¹	arwəʔat ¹	arbat ¹	arβat ¹	arbat ¹	arβat ¹	arbat ¹	arbat ¹	arbat ¹	arat ¹
76	fox	kəborə ¹	deegallo	kəwəra ¹	walake ²	kəwəra ¹	kəwəra ¹	kəwəra ¹	ing'əragiya ⁴	səwəndia ⁴	yəg'ərag'ə ⁴	kəwəra ¹	zəgədo ¹
77	friday	arb	gimma ¹	gimat ¹	gimʔat ¹	addara ²	gimat ¹	gimat ¹	gimʔat ¹	addara ²	addara ²	addara ²	gimat ¹
78	frog	inukurari ²	raaččə	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	kəwəčə ¹	inukurari ²
79	full	mulu ¹	gutuu	mura ¹	mura ¹	mura ¹	mura ¹	mura ¹	mura ¹	mura ¹	mura ¹	mura ¹	mura ¹
80	garlic	naččə šinkurt	kullubi	tumma ¹	tumma ¹	tumma ¹	tumma ¹	tumma ¹	tumma ¹	tumma ¹	tumma ¹	tumma ¹	tumma ¹
81	gazelle	midaččə	kurup p'e	gimbe ¹	gimba ¹	gimbe ¹	g'imbe ¹	gimba ¹	gimba ¹	g'imbe ¹	g'imbe ¹	g'imbe ¹	gedemo ²
82	giraffe	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	kəččine ¹	sologe ²
83	give up	təwə ¹	diše	wəte ¹	itaʔut ²	wəte ¹	wəte ¹	wəte ¹	ətaʔut ²	wəgəfir ³	wəgəfir ³	wəte ¹	idagot ⁴
84	goat	fiyyal ¹	reʔe	fek ¹	fəŋəŋə ²	fiyək ¹	fek ¹	fek ¹	fəŋə ²	fiyyal ¹	fek ¹	fəŋəŋə ²	fek ¹
85	griddle	miʔad ¹	kibaa	midad ¹	miʔad ¹	midad ¹	midad ¹	midad ¹	midad ¹	miʔad ¹	midad ¹	midad ¹	miʔad ¹
86	guava	zəytun ¹	zeeytuna ¹	zəytun ¹	zəytuna ¹	zəytun ¹	zəytuna ¹	zəytun ¹	zəytun ¹	zəytun ¹	zəytun ¹	zəytun ¹	zəytun ¹
87	had	nəbbəwə ²	kəbature	b'anən ¹	banənda ¹	b'anən ¹	banənda ¹	banənda ¹	banənda ¹	nəbbəwə ²	b'annən ¹	banənnəmitt ¹	nəy ²
88	hair	šəgur	rifeensa	gunər ¹	gun'ər ¹	gunnər ¹	gunər ¹	gunər ¹	gun'ər ¹	gunnən ¹	gunnən ¹	gunnən ¹	dum ²
89	handkerchief	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹	məhərrəb ¹
90	hard (adv)	bəhayl ¹	haalaan	banəndəg ¹	bəndəg ¹	yənnənəm ²	bəŋona ⁴	bəŋona ⁴	gerika ³	bəhayl ¹	bəhayl ¹	bəhayl ¹	bəyadidnə ⁶
91	hat	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹	kəfiyya ¹
92	he	issu	isa	hut ¹	huda ¹	hut ¹	xut ¹	hut ¹	huda ¹	k'a ¹	hut ¹	x'a ¹	uha ¹
93	head	ras ²	mataa	gunər ¹	iras ²	gunnər ¹	gunər ¹	gunər ¹	gun'ər ¹	gunnən ¹	gunnən ¹	gunnən ¹	dum ²
94	headaddress	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	šaš ¹	gufi ³
95	heart	libb ³	onne ²	xin ¹	xin ¹	xinn ¹	x'in ¹	xin ¹	ana ²	libb ³	hin ¹	libb ³	wəzəna ⁴
96	heifer	gidər	waatiyyo	anɜ ¹	anɜ ¹	anɜ ¹	anɜ ¹	anɜ ¹	anɜ ¹	zəbay ²	anɜ ¹	g'əbəz ²	zəba ²
97	helped	aggəzo ¹	gargaare	agəzəm ¹	eporə ²	aggəzəm ¹	agəzəm ¹	agəzəm ¹	eporə ²	orəddam ³	aggəzo ¹	orəddam ³	agəzo ¹
98	hen	doro	handaakko ³	kutara ¹	ungoro ²	kutara ¹	kutara ¹	kutara ¹	gəra ²	kuttəna ¹	gəra ²	kuttəna ¹	inčako ³
99	him	issun	isa	yəhut ¹	huda ¹	huta ¹	xut ¹	huta ¹	huda ¹	yək'a ¹	yəhut ¹	yək'a ¹	uhaana ¹
100	hippopotamus	gumarre ¹	gumaarree ¹	g'əmare ¹	gəwira ²	gumarre ¹	gumar'ə ¹	gumarre ¹	gowana ³	gomarre ¹	gumarre ¹	g'əmare ¹	gomare ¹
101	his	yəssu	kan isaa	yəhut ¹	yəxut ¹	yəhut ¹	yəxut ¹	yəhut ¹	ihuda ¹	yək'a ¹	yəhut ¹	yək'a ¹	yuha ¹
102	horse	foras ¹	farda ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹	fəroz ¹
103	hot	muk ¹	hoʔaa	yəmwəka ¹	mwok ¹	mwok ¹	mwok ¹	mwok ¹	mwok ¹	yəmwəka ¹	yəmwəka ¹	m'əʔ ¹	muk ¹
104	hour	soʔat ¹	saʔaa ¹	sat ¹	saʔat ¹	sat ¹	sat ¹	soʔat ¹	saʔat ¹	sat ¹	sat ¹	sat ¹	sat ¹
105	how	indet	akkamitti	namir ¹	etəho ²	namir ¹	namir ¹	otəka ²	namkom ³	namir ¹	namir ¹	namir ¹	aynako ⁴
106	hundred	məto ²	clibba	bəkir ¹	bəʔər ¹	bəkir ¹	məto ²	bəkir ¹	bəʔər ¹	məto ²	məto ²	bəʔər ¹	bəkir ¹
107	hyena	gib	waraabessa ²	g'əməčə ¹	gončə ¹	g'əməčə ¹	g'əməčə ¹	gončə ¹	gončə ¹	g'əməčə ¹	gončə ^{1</}		

181	song	zəfən	sirba	lale ¹	das ²	lale ¹	weg ³	lale ¹	lale ¹	dars ²	dars ²	dərris ⁴	šibot ⁵
182	soon	wədiyaw ⁴	batahumatti	hintołm ¹	hangizzi ²	hingiz ² e ²	ximtoxim ¹	kinəga ²	himaga ²	zagizze ²	kəmo ³	wədiyaw ⁴	wəktəy ⁵
183	sorghum	mašilla ¹	bišingaa	mašilla ¹	mašulla ¹	mašilla ¹	mašilla ¹	mašilla ¹	m ¹ asina ¹	mašilla ¹	mašilla ¹	mašilla ¹	mašulle ¹
184	steer	wəyən ¹	waatiyyo	wəfer ¹	wəfir ¹	wəfir ¹	wəfer ¹	wəfir ¹	wəfir ¹	mofen ¹	wəfir ¹	wəfen ¹	mofən ¹
185	stew	wət ¹	itto	wət ¹	wət ¹	wət ¹	wət ¹	wət ¹	wət ¹	wət ¹	wət ¹	wət ¹	wət ¹
186	stirring rod	mamasəya ²	kokorsaa	wəgəŋga ¹	mammase ²	wəgəŋga ¹	wəgəŋga ¹	wəgəŋga ¹	wəgəŋga ¹	wəgəŋga ¹	wəgəŋga ¹	wəgəŋga ¹	hunkutma ³
187	stone	dəŋgay	dagaa	imur ¹	əwru ²	imur ¹	imur ¹	imur ¹	imur ¹	imur ¹	imur ¹	imur ¹	um ²
188	stool	bərcuma ¹	bərcuma ¹	b ¹ ərcuma ¹	bərcumma ¹	b ¹ ərcuma ¹	bərcuma ¹	wərcuma ¹	bərcima ¹	wənbər ²	bərcuma ¹	b ¹ ərcimma ¹	katta ³
189	stronger	fənkarra ⁴	irra címaa	ğədanə ¹	tuʔə ²	tiwwo ²	tuwə ²	tiwwo ²	tuʔə ²	bərtu ³	bərtu ³	fənkarra ⁴	ğədit ¹
190	sucked	fəbba ¹	hüge	təp ¹ əm ¹	fəppə ¹	fəbb ¹ əm ¹	fəpp ¹ əm ¹	təp ¹ əm ¹	fəp ¹ ə ¹	fəbbəm ¹	fəbba ¹	fəbb ¹ əm ¹	tobe ¹
191	suddenly	dingət ²	tasa	mmatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹	immatəg ¹
192	sun	šəhay	adun ²	čet ¹	ayed ²	čet ¹	čet ¹	čet ¹	iwəya ¹	yimur ²	čet ¹	aret ²	ayir ²
193	sunday	ihud	dilbata	wursənbət ¹	abba ²	wursənbət ¹	wursənbət ¹	wursənbət ¹	wursənbət ¹	wursənbət ¹	wursənbət ¹	wursənbət ¹	gidirsənbət ¹
194	sunflower	suf ¹	sunfi ¹	suf ¹	suf ¹	suf ¹	suf ¹	suf ¹	suf ¹	suf ¹	suf ¹	suf ¹	suf ¹
195	sweat	lab	hukkaa	wizat ¹	azat ¹	wizat ¹	wizat ¹	lab ²	azat ¹	wizat ¹	wizat ¹	wizat ¹	wizat ¹
196	takeoff	awəlləka ²	baase	watməŋig ¹	awənaga ²	atfəm ¹	awtə ¹	wawta ¹	atwanukt ¹	awəlləka ²	wawlək ²	awəlləka ²	šeeləka ²
197	tangerine	məndarin ¹	mandariina ¹	məndarəni ¹	məndarin ¹	məndarəni ¹	məndarin ¹	məndarəni ¹	məndarəni ¹	məndarin ¹	məndarin ¹	məndarəni ¹	məndarəni ¹
198	teff	fəf ¹	fəaifi ¹	fəfi ¹	fəf ¹	fəfi ¹	fəfi ¹	fəfi ¹	fəfi ¹	fəfi ¹	fəfi ¹	fəfi ¹	fəfi ¹
199	ten	assir ¹	kudān	assir ¹	assir ¹	assir ¹	assir ¹	assir ¹	assir ¹	assir ¹	assir ¹	assir ¹	assir ¹
200	that	ya ²	san	xix ¹	xada ¹	za ²	xix ¹	hata ¹	xada ¹	za ²	za ²	za ²	hatty ¹
201	then	kəziya	sanbooda	təhanke ¹	tohaʔegad ¹	təhanke ¹	taxanke ¹	təhʔang ¹ e ¹	təhim ¹	təh ¹ a ²	bohi ³	təhimange ⁴	toʔizof ⁵
202	they	innəssu	isaan	xino ¹	huno ¹	xino ¹	xino ¹	xinu ¹	xinowa ¹	kinnəm ¹	xinno ¹	xinnaw ¹	uxun ¹
203	thick	wəfram	furdaa	ğəŋgir ¹	gušir ¹	ğəŋgir ¹	ğəŋgir ¹	ğəŋgir ¹	ğəŋgir ¹	ğəŋgir ¹	ğəŋgir ¹	ğəŋgir ¹	ğəŋgir ¹
204	thief	leba ¹	hattuu	neba ¹	fəngaya ²	səwəhə ³	nəba ¹	neba ¹	fəngaya ²	leba ¹	leba ¹	leba ¹	rangā ¹
205	thin	kəccin ²	kallaa	sisə ¹	kəccin ²	sisə ¹	sisə ¹	sisə ¹	kəccin ²	asfətib ³	kəccin ²	səssa ¹	kəccin ²
206	this	yih ¹	kun	zi ¹	wa ²	zix ¹	zi ¹	zi ¹	wəda ²	zi ¹	zi ¹	zi ¹	innə ³
207	three	sost ¹	sadi ¹	sost ¹	sost ¹	sost ¹	sost ¹	sost ¹	sost ¹	sost ¹	sost ¹	sost ¹	šest ¹
208	thursday	hamus ¹	kamisa ¹	amus ¹	amus ¹	amus ¹	am ¹ is ¹	amus ¹	amus ¹	xamus ¹	amus ¹	amus ¹	xamus ¹
209	tiger	nəbir ³	kətreensa	ğəg ¹ ara ¹	zəgə ¹	zəg ¹ ara ¹	zəg ¹ ara ¹	zəg ¹ ara ¹	zəgə ¹	buswa ²	zəgə ¹	zəgə ¹	naaw ³
210	till	dirəs ¹	haanga	hihedər ¹	haʔed ²	dar ¹	dar ¹	dar ¹	sənəga ³	dirəs ¹	sinn ⁴	dirəs ¹	gaango ⁵
211	time	gizze ¹	yaroo	gidad ²	əga ²	gam ¹ ə ¹	gizze ¹	gizze ¹	gizze ¹	gizze ¹	gizze ¹	gizze ¹	wəkt ¹
212	to	lə ³	f	ya ¹	ʔə ²	lə ³	ya ¹	ya ¹	ʔə ²	lə ³	ya ¹	ya ¹	lə ³
213	today	zare	harra	ək ¹ a ¹	akka ¹	ək ¹ a ¹	ək ¹ a ¹	ək ¹ a ¹	ək ¹ a ¹	axum ¹	ək ¹ a ¹	ək ¹ a ¹	awge ²
214	tomato	tumatin ¹	timaatiim ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹	tumatin ¹
215	tomorrow	nəga ¹	boru	nəga ¹	nəga ¹	nəga ¹	nəga ¹	nəga ¹	nəga ¹	nəga ¹	nəga ¹	nəga ¹	gees ²
216	too	bəfəm ³	baay ¹ ee	niškar ¹	haməʔad ²	niškar ¹	niškar ¹	niškar ¹	bušeʔəha ²	burim ³	bəfəm ³	niškar ¹	siwəm ⁵
217	traveler	məngədañña ³	emala ¹	eməna ¹	meñña ¹	eməna ¹	eməna ¹	eməna ¹	eməna ¹	əngədañña ³	imətañña ¹	məngədañña ³	ungawəro ⁴
218	trivet	gulličča ²	sumsuma	gonziya ¹	gamziye ¹	gamziya ¹	gulličča ²	g ¹ amziyye ¹	midačča ²	gawziya ¹	akufa ³	gonziya ¹	siškar ¹
219	trouser	surri ²	kofo	bolale ¹	surri ²	sulle ²	bolale ¹	bolale ¹	bolale ¹	luko ³	surri ²	surri ²	bolale ¹
220	two	hulətt ²	lama	x ¹ et ¹	hur ¹ ət ¹	x ¹ ett ¹	x ¹ et ¹	x ¹ et ¹	wir ¹ et ¹	kətt ¹	x ¹ et ¹	x ¹ ett ¹	ost ²
221	upon	šayulay	gubbaarra	bəforəta ¹	bəforfor ¹	bəforəta ¹	bəforəta ¹	bəforəta ¹	bəforəta ¹	bəforəta ¹	bəforəta ¹	bəforəta ¹	bəforəta ¹
222	us	ihñan ¹	nu	yina ¹	imə ¹	yina ¹	yina ¹	ina ¹	ina ¹	ihñan ¹	inna ¹	ihñan ¹	yəñna ¹
223	waist-band	məkənət	sabbata	azgat ¹	azgəttə ¹	azgard ¹	azgard ¹	azgard ¹	azgard ¹	dəkot ²	azlagidd ¹	azgalidd ¹	itto ³
224	war	fər	waraana	gar ¹	ariv ²	arib ²	gar ¹	gar ¹	gar ¹	gar ¹	gar ¹	gar ¹	arib ²
225	was	nəbbar ²	ture	bəna ¹	banəda ¹	baanna ¹	bəna ¹	bəna ¹	banəda ¹	nəbbəra ¹	banna ¹	banna ¹	narə ²
226	water	wiha	bišaan	ixa ¹	ixə ¹	ixa ¹	ixa ¹	ixa ¹	ixa ¹	yiga ¹	əga ¹	iga ¹	məy ²
227	wednesday	ərob ¹	roobi ¹	əro ¹	arbe ¹	ərow ¹	əro ¹	əro ¹	əro ¹	xərob ¹	ərob ¹	ərob ¹	harbe ¹
228	were	nəbbaru ²	turan	banəwo ¹	banowəttə ¹	banəbo ¹	banəbo ¹	banəwo ¹	banowəttə ¹	nəbbərim ²	banna ¹	banimutt ¹	narū ²
229	what	min ¹	maal ¹	mirkar ¹	mirafor ¹	məʔe ¹	mirkar ¹	mirkar ¹	mirkar ¹	min ¹	mirkar ¹	mirkar ¹	min ¹
230	wheat	sinde ¹	kamadi	sina ¹	sine ²	sime ¹	sina ¹	sine ¹	sina ¹	sirre ¹	sirre ¹	sirre ¹	sirə ¹
231	which	yətiñhaw	kam	əndehuta ²	ətəda ¹	indeyətə ¹	etətə ¹	etətə ¹	ətəda ¹	yitta ¹	etti ¹	infewe ¹	aytay ¹
232	who	man ¹	eeñnu	m ¹ an ¹	man ¹	m ¹ an ¹	m ¹ an ¹	m ¹ an ¹	m ¹ an ¹	ma ¹	m ¹ a ¹	ma ¹	maani ¹
233	whoever	maniñhawm ²	kamiyyu	m ¹ animənda ²	attimanir ¹	indeyənda ²	etətə ¹	etənda ¹	itindra ¹	yittañna ¹	ettənnə ¹	mannimima ²	aynaanna ³
234	wide (adj)	səffi	balʔa	bəti ¹	bəttad ¹	bəttit ¹	bəti ¹	bəti ¹	bəti ¹	bəti ¹	bəttit ¹	bəttit ¹	fətt ¹
235	wind	nifas ¹	killeensa	infəs ¹	infəs ¹	nəfas ¹	nifas ¹	infəs ¹	nəfas ¹	nəfas ¹	nifas ¹	infəs ¹	nifas ¹
236	with	gar ¹	wagğūn	təhur ²	gi ¹	gamwa ¹	gam ¹ a ¹	gəz ¹ a ¹	goy ¹	təg ¹	ge ¹	ge ¹	təhuhə ²
237	woman	setiyyo	gaarti	mišt ¹	mišt ¹	mišt ¹	mišt ¹	mišt ¹	mišt ¹	mišt ¹	mišt ¹	mišt ¹	mišt ¹
238	word	kāl ¹	gačča	kār ¹	kāl ¹	kār ¹	kār ¹	kār ¹	kār ¹	kāl ¹	kāl ¹	kāl ¹	kāl ¹
239	wrapped (adj)	fəkūll ¹	maramaa	fəkūll ¹	dububūl ²	fəkūll ¹	fəkūll ¹	fəkūll ¹	fəkūll ¹	fəkūll ¹	fəkūll ¹	fəkūll ¹	fəkūll ¹
240	you (2.M.SG)	antə ¹	ati ¹	axə ¹	axə ¹	zəhə ¹	axə ¹	axə ¹	axə ¹	dəxə ¹	axə ¹	ata ²	ata ²

B.2. List of words for Word Categorization test

The following list of words was used in the word categorization test to measure mutual intelligibility and to determine the functional distance among the selected language varieties.

cloths	Body parts	Kitchen utilities	Fruits	Food type
Shoes	Finger	Spoon	Banana	Bread
Shirt	Lip	Ladle	Mango	'Kocho'
Hat	Eye	Pan	Orange	'Injera'
Belt	Arm	Knife	Berry	Stew
Trouser	Breast	Cutting board	Guava	Pancake
Handkerchief	Leg	Griddle	Cherimoya	Roasted meat
Dress	Chest	Stirring rod	Coke	Mush
Shorts	Eye	Kettle	Tangerine	'Besso'
Waist-band	Hair	Food-table	Lemon	Porridge
Headdress	Neck	Plate	Doviyalis abyssinica	Roasted grain

Domestic animal	Furniture	Vegetables	Wild animal	Cereals
Hen	Table	Cabbage	Elephant	Barely
Ox	Chair	Pepper	Lion	Wheat
Camel	Shelf	Tomato	Tiger	Maize
Donkey	Locker	Onion	Hyena	Pea
Goat	Bed	Potato	Crocodile	Fava bean
Sheep	Sofa	Carrot	Giraffe	Sorghum
Dog	Stool	Garlic	Monkey	'Teff'
Cat	Chassis	Pumpkin	Ape	Bean
Horse	Mirror	Sweet potato	Fox	Lentil
Mule	Box	Basil	Gazelle	Chickpea

Appendix C. Additional results

C.1. Results of fuzzy clustering

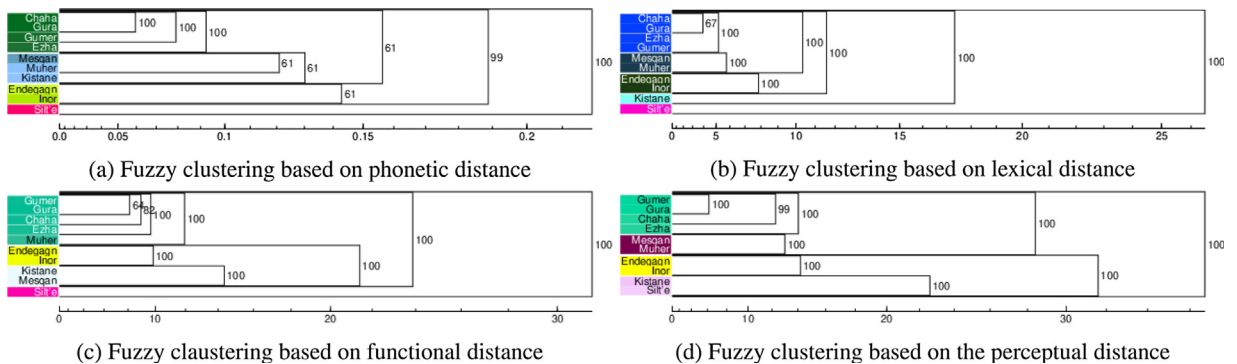


Fig. 11. Fuzzy clustering based on the structural, functional and perceptual distances.

C.2. Multidimensional scaling-second dimension

866 C.2 Multidimensional scaling-second dimension

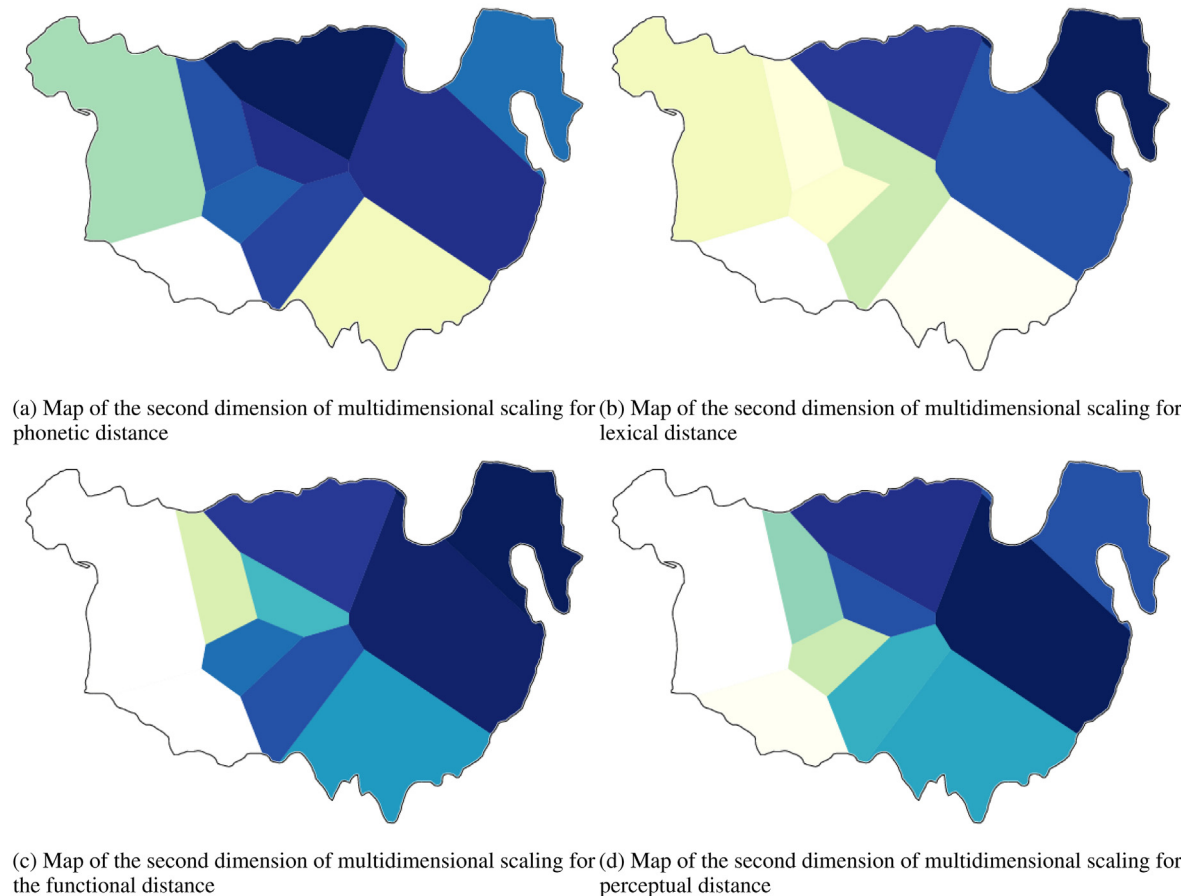


Fig. 12. Map of the second dimension of multidimensional scaling for the structural, functional and perceptual distances.

References

- Abu-Rabia, S., 1996. Attitudes and cultural background and their relationship to reading comprehension in a second language: a comparison of three different social contexts. *Int. J. Appl. Linguist.* 6 (1), 81–105.
- Abu-Rabia, S., 1998. The influence of the Israel–Arab conflict on Israeli–Jewish students learning Arabic as a third language. *Lang. Cult. and Curricul.* 11 (2), 154–164.
- Ahland, M., 2003. Interlectal intelligibility between Gurage speech varieties. In: *North American Conference on Afroasiatic linguistics*, pp. 45–71.
- Anthony, L., 2004. Antconc: a learner and classroom friendly, multi-platform corpus analysis toolkit. In: *Proceedings of IWLLeL*, pp. 7–13.
- Beijering, K., Gooskens, C., Heeringa, W., 2008. Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. *Linguist. Netherlands* 15, 13–24.
- Bender, M., 1971. The languages of Ethiopia: a new lexicostatistic classification and some problems of diffusion. *Anthropol. Linguist.* 13 (5), 165–288.
- Bender, M., Cooper, R.L., Ferguson, C., 1972. Language in Ethiopia: implications of a survey for sociolinguistic theory and method. *Lang. Soc.* 1 (2), 215–233.
- Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T.O., Basbøll, H., 2008. Early vocabulary development in Danish and other languages: a CDI-based comparison. *J. Child Lang.* 35 (3), 619–650.

- Bucholtz, M., Bermudez, N., Fung, V., Edwards, L., Vargas, R., 2007. Hella Nor Cal or totally So Cal? The perceptual dialectology of California. *J. English Linguist.* 35 (4), 325–352.
- Casad, E., 1987. *Dialect Intelligibility Test*. Summer Institute of Linguistics, Dallas, Texas.
- Chiswick, B., Miller, P.W., 2005. Linguistic distance: a quantitative measure of the distance between English and other languages. *J. Multiling. Multicult. Dev.* 26 (1), 1–11.
- Demeke, G., 2001. The Ethio-Semitic languages (re-examining the classification). *J. Ethiopian Stud.* 34 (2), 57–93.
- ENSA, 2007. *Ethiopian Statistical Report of the 2007 Population and Housing Census: Population Size by Age and Sex, South Nation Nationalities*. Federal Democratic Republic of Ethiopia, Population Census Commission, Addis Ababa.
- Faber, A., 1997. Genetic subgrouping of the Semitic languages. *Semitic Lang.* 15, 3–15.
- Fleming, H., 1968. Ethiopic language history: testing linguistic hypotheses in an archaeological and documentary context. *Ethnohistory* 15 (4), 353–388.
- Goebel, H., 2010. Dialectometry and quantitative mapping. In: Lameli, A., Kehrein, R., Rabanus, S. (Eds.), *Language and Space. An International Handbook of Linguistic Variation*, Vol. 2. De Gruyter Mouton, Berlin, New York, pp. 433–457.
- Goeman, A., 1999. Dialects and subjective judgement of speakers. In: Daniel, L., Preston, D. (Eds.), *Handbook of Perceptual Dialectology*, Vol. 2. John Benjamins, Amsterdam, pp. 135–136.
- Goldenberg, G., 1977. The Semitic languages of Ethiopia and their classification. *Bull. School Orient. Afr. Stud.* 40 (3), 461–507.
- Golubović, J., Sokolić, N., 2013. Their language sounds aggressive: a matched guise study with Serbian and Croatian. In: Gooskens, C., Van Bezooijen, R. (Eds.), *Phonetics in Europe. Perception and Production*. Peter Lang, Frankfurt, pp. 35–57.
- Gooskens, C., 2005. How well can Norwegians identify their dialects? *Nord. J. Linguist.* 28 (1), 37–60.
- Gooskens, C., 2007. The contribution of linguistic factors to the intelligibility of closely related languages. *J. Multiling. Multicult. Dev.* 28 (6), 445–467.
- Gooskens, C., 2013. Experimental methods of measuring intelligibility of closely related language varieties. In: Robert, B., Richard, C., Ceil, L. (Eds.), *The Oxford Handbook of Sociolinguistics*. Oxford University Press, Oxford, pp. 195–213.
- Gooskens, C., 2018. Dialect intelligibility. In: Boberg, C., Nerbonne, J., Watt, D. (Eds.), *Handbook of Dialectology*. Wiley-Blackwell, Hoboken, NJ, pp. 204–218.
- Gooskens, C., Heeringa, W., 2004. Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Lang. Variat. Change* 16 (3), 189–207.
- Gooskens, C., Heuven, V., 2018. How well can intelligibility of closely related languages in Europe be predicted by linguistic and non-linguistic variables? *Linguist. Approach. Biling.* 4 (8), 1–29.
- Gooskens, C., Van Heuven, V.J., Van Bezooijen, R., Pacilly, J.J., 2010. Is spoken Danish less intelligible than Swedish? *Speech Commun.* 52 (11–12), 1–33.
- Grimes, J., 1990. Interpreting sample variation in intelligibility tests. In: Williams, G. (Ed.), *Synchronic and Diachronic Approaches to Linguistic Variation and Change*. Georgetown University, Washington, DC, pp. 138–146.
- Grimes, J., 1995. *Language Survey Reference Guide*. Summer Inst of Linguistics, Dalla.
- Gutt, E.A., 1980. Intelligibility and interlingual comprehension among selected Gurage speech varieties. *J. Ethiopian Stud.* 14, 57–85.
- Heeringa, W., 2004. *Measuring Dialect Pronunciation Differences Using Levenshtein Distance* University of Groningen (Unpublished doctoral dissertation).
- Heeringa, W., Kleiweg, P., Gooskens, C., Nerbonne, J., 2006. Evaluation of string distance algorithms for dialectology. In: *Proceedings of the Workshop on Linguistic Distances*. Association for Computational Linguistics, Sydney, pp. 51–62.
- Hetzron, R., 1972. *Ethiopian Semitic: Studies in Classification* (No. 2). Manchester University Press.
- Hetzron, R., 1977. *The Gunn“an-Gurage Languages*, Vol. 12. Istituto Universitario Orientale, Naples.
- Hetzron, R., Bender, M., 1976. The Ethiosemitic languages. In: Bender, L. (Ed.), *Languages in Ethiopia*. Oxford University Press, London, pp. 2–29.
- Hudson, G., 2000. Ethiopian semitic overview. *J. Ethiopian Stud.* 33 (2), 75–86.
- Hudson, G., 2013. *Northeast African Semitic: Lexical Comparisons and Analysis*. Harrassowitz, Wiesbaden.
- Kessler, B., 1995. Computational dialectology in Irish Gaelic. In: *Proceedings of the Seventh Conference on European Chapter of Association for Computational Linguistics*. Morgan Kaufmann, Burlington, pp. 60–66.
- Labov, W., 2001. *Principles of Language Change: Social Factors*. Blackwell, Malden, MA.
- Leinonen, T., Ç”oltekin, Ç., Nerbonne, J., 2016. Using Gabmap. *Lingua* 178, 71–83.
- Leslau, W., 1969. Toward a classification of the Gurage dialects. *J. Semitic Stud.* 14 (1), 96–109.
- Leslau, W., 1979. *Etymological Dictionary of Gurage (Ethiopic): English-Gurage Index*, Vol. 2. Otto Harrassowitz Verlag.
- Leslau, W., 1992. *Gurage Studies: Collected Articles*. Otto Harrassowitz Verlag.
- Menuta, F., 2015. *Intergroup Communication Among Gurage: A Study in Intelligibility, Inter-Lingual Comprehension and Accommodation* Addis Ababa University (Unpublished doctoral dissertation).
- Meyer, R., 2011. Gurage. In: Weninger, S. (Ed.), *The Semitic Languages: An International Handbook*. DeGruyter Mouton, Berlin, pp. 1220–1257.
- Meyer, R., 2014. Finiteness in Gurage languages. In: Amha, A., Meyer, R., Yvonne, T. (Eds.), *Explorations in Ethiopian Linguistics: Complex Predicates, Finiteness and Interrogativity*. Harrassowitz, Wiesbaden, pp. 225–258.
- Montgomery, C., 2007. *Northern English Dialects: A Perceptual Approach* University of Sheffield (Unpublished doctoral dissertation).
- Nerbonne, J., 2010. Measuring the diffusion of linguistic change. *Philos. Trans. R. Soc. London: Biol. Sci.* 365 (1559), 3821–3828.
- Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P., Leinonen, T., 2011. Gabmap – a web application for dialectology. *Dialectologia: Rev. Electrón.* 2 (2), 65–89.
- Nerbonne, J., Heeringa, W., 2001. Computational comparison and classification of dialects. *Dialectol. Geolinguist.* 9, 69–83.
- Niedzielski, N., 1999. The effect of social information on the perception of sociolinguistic variables. *J. Lang. Social Psychol.* 18 (1), 62–85.
- Pavlenko, A., 2006. Russian as a lingua franca. *Annu. Rev. Appl. Linguist.* 26, 78–99.
- Pearce, M., 2009. A perceptual dialect map of North East England. *J. English Linguist.* 37 (2), 162–192.

- Plichta, B., Preston, D., 2005. The /ay/s have it the perception of /ay/as a north-south stereotype in United States English. *Acta Linguist. Hafniensia* 37 (1), 107–130.
- Preston, D., 1996. Where the worst English is spoken. In: Schneider, E. (Ed.), *Focus on USA*. John Benjamins, Amsterdam, pp. 297–361.
- Preston, D., 2010. Mapping the geolinguistic space of the brain. In: Lameli, A., Kehrein, R., Rabanus, S. (Eds.), *Language and Space. An International Handbook of Linguistic Variation: Language Mapping*, Vol. 2. De Gruyter Mouton, Berlin, New York, pp. 121–141.
- Prokić, J., Nerbonne, J., et al., 2013. Analyzing dialects biologically. In: Heiner, F., Hans, G., Thoresten, H., William, M. (Eds.), *Classification and Evolution in Biology, Linguistics and the History of Science*. Franz Steiner, Stuttgart, pp. 147–159.
- Snoek, C., 2014. Review of Gabmap: doing dialect analysis on the web. *Lang. Document. Conserv.* 8, 192–208.
- Swadesh, M., 1950. Salish internal relationships. *Int. J. Am. Linguist.* 16 (4), 157–167.
- Swadesh, M., 1955. Towards greater accuracy in lexicostatistic dating. *Int. J. Am. Linguist.* 21 (2), 121–137.
- Tamasi, S., 2003. *Cognitive Patterns of Linguistic Perceptions* University of Georgia, Georgia (Unpublished doctoral dissertation).
- Tang, C., van Heuven, V., 2009. Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119 (5), 709–732.
- Tang, C., van Heuven, V.J., 2007. Mutual intelligibility and similarity of Chinese dialects: predicting judgments from objective measures. *Linguist. Netherlands* 24 (1), 223–234.
- Tang, C., Van Heuven, V.J., 2015. Predicting mutual intelligibility of Chinese dialects from multiple objective linguistic distance measures. *Linguistics* 53 (2), 285–312.
- Van Bezooijen, R., Gooskens, C., 2007. Interlingual text comprehension: linguistic and extralinguistic determinants. In: Jan, T., Ludger, Z. (Eds.), *Receptive Receptive Multilingualism, Linguistic Analyses, Language Policies And Dialectic Concepts*. John Benjamins, Amsterdam, pp. 249–263.
- Voegelin, C.F., Harris, Z.S., 1951. Methods for determining intelligibility among dialects of natural languages. *Proc. Am. Philos. Soc.* 95 (3), 322–329.